

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 00:44:53 ; Search time 164 Seconds
(without alignments)
14.150 Million cell updates/sec

Title: US-10-027-015A-4

Perfect score: 32
Sequence: 1 IGDIGD 6

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 50 summaries

Database: A_Geneseq_16Dec04:*

1: Geneseqp1980s:*\n2: Geneseqp1990s:*\n3: Geneseqp2000s:*\n4: Geneseqp2001s:*\n5: Geneseqp2002s:*\n6: Geneseqp2003as:*\n7: Geneseqp2003bs:*\n8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	6	ADG43179	Adg43179 IGD-motif
2	32	100.0	6	ADG43178	Adg43178 IGD-motif
3	32	100.0	443	ABP65456	Abp65456 Bifidobac
4	32	100.0	515	ADS29151	Ads29151 Bacterial
5	31	96.9	59	AAW22623	AAW22623 Type II t
6	31	96.9	69	AAW22623	AAW22623 Human sec
7	31	96.9	81	ABO55461	ABO55461 Human gen
8	31	96.9	86	AAW22623	AAW22623 Human sec
9	31	96.9	129	AAU18547	AAU18547 Human cyt
10	31	96.9	149	AAU51226	AAU51226 Propionib
11	31	96.9	149	ABM47745	ABM47745 Propionib
12	31	96.9	156	AAU18521	AAU18521 Human cyt
13	31	96.9	166	AAW78545	AAW78545 Human pro
14	31	96.9	166	ABR41956	ABR41956 Human cot
15	31	96.9	166	ADP75248	ADP75248 Mouse cot
16	31	96.9	166	ADP75248	ADP75248 Prostate
17	31	96.9	166	ADJ70104	ADJ70104 Human hea
18	31	96.9	166	ADJ24536	ADJ24536 Human mod
19	31	96.9	173	AAU52312	AAU52312 Propionib
20	31	96.9	173	ABM48831	ABM48831 Propionib
21	31	96.9	201	ADQ10039	Adq10039 B.afze111
22	31	96.9	205	AAU18546	AAU18546 Human cyt
23	31	96.9	220	AAW79529	AAW79529 Human pro
24	31	96.9	220	ADQ10053	Adq10053 B.afze111
25	31	96.9	342	ADJ48363	Adj48363 Mize oil
			364	ADG45014	ADG45014 Bacterial

26	31	96.9	693	ADS22843	ADS22843 Bacterial
27	31	96.9	1971	ABP22942	ABP22942 Mouse GAN
28	31	96.9	1971	ABP22942	ABP22942 Murine SH
29	31	96.9	1971	ABP22942	ABP22942 Murine CA
30	31	96.9	1971	ADQ31144	Adq31144 Mouse ger
31	30	93.8	58	AAU65282	AAU65282 Propionib
32	30	93.8	58	ABM61801	ABM61801 Propionib
33	30	93.8	78	AAU59294	AAU59294 Propionib
34	30	93.8	78	ABM55813	ABM55813 Propionib
35	30	93.8	87	ADM27005	ADM27005 Hyperther
36	30	93.8	122	ADP43918	ADP43918 Chicken w
37	30	93.8	122	ADP43918	ADP43918 Propionib
38	30	93.8	127	ABM59889	ABM59889 Propionib
39	30	93.8	127	ABM65515	ABM65515 Propionib
40	30	93.8	261	AAU65828	AAU65828 Propionib
41	30	93.8	261	ABM62347	ABM62347 Propionib
42	30	93.8	275	ABG14312	ABG14312 Novel hum
43	30	93.8	302	ABP77414	ABP77414 Fungal me
44	30	93.8	308	ABP66144	ABP66144 Bifidobac
45	30	93.8	353	ABM65924	ABM65924 Propionib
46	30	93.8	456	ABO71148	ABO71148 Pseudomon
47	30	93.8	500	ADQ36971	Adq36971 Cell prol
48	30	93.8	505	ADQ15641	Adq15641 Rice stre
49	30	93.8	544	ADQ15026	Adq15026 Synchococ
50	30	93.8	544	ADQ15026	ADQ15026 Klebsiell

ALIGNMENTS

RESULT 1	ADG43179	standard; peptide; 6 AA.
XX	ADG43179	
XX	ADG43179	
AC	ADG43179	
XX	ADG43179	
DT	26-FEB-2004	(first entry)
XX	IGD-motif peptide #SEQ ID 4.	
DE	Angiogenic composition; isoleucine-glycine-aspartic acid; IGD;	
XX	angiogenic growth factor; cell migration; angiogenesis; ischemia; heart;	
KW	polyvinyl pyrrolidone; cyclic.	
KW	Unidentified.	
OS	WO2003059436-A2.	
XX	24-JUL-2003.	
PD	23-DEC-2002; 2002MO-US041484.	
XX	21-DEC-2001; 2001US-00027015.	
XX	(CENT-) CENTERPULSE BIOLOGICS INC.	
XX	Akella R, Ranieri J;	
XX	WPI; 2003-697294/66.	
DR	Angiogenic composition comprising peptides containing isoleucine-glycine-	
PT	aspartic acid that stimulate cell migration, and angiogenic growth	
PT	factor useful for promoting myocardial or peripheral angiogenesis.	
XX	Claim 1, SEQ ID NO 4; 22pp; English.	
XX	The invention relates to an angiogenic composition (I) comprising at	
CC	least one peptide chosen from a group of peptides containing isoleucine-	
CC	glycine-aspartic acid (IGD-containing peptide), and at least one	
CC	angiogenic growth factor other than the at least one peptide, where the	
CC	peptide stimulates cell migration. Also disclosed is a composition (III)	
CC	that is active for promoting cell migration and/or angiogenesis under	
CC	cell growth promoting conditions comprising an IGD-containing peptide,	

CC and a matrix material. Compositions of the invention are useful for
CC promoting myocardial angiogenesis which involves administering (I) or
CC (III) intramyocardially to an ischaemic area of the heart of an
CC individual, to enhance vascular endothelial cell migration and/or
CC proliferation. They are also useful for promoting peripheral angiogenesis
CC in an ischaemic region of an organ or tissue fed by a peripheral vessel,
CC by enhancing vascular endothelial cell migration and/or proliferation at
CC the ischaemic region, or for enhancing blood flow to an ischaemic tissue
CC of the body, where the composition stimulates vascular endothelial cell
CC migration and/or proliferation sufficient to restore or increase blood
CC flow to the ischaemic tissue. The physiologically acceptable carrier in
CC the composition is polyvinyl pyrrolidone. The current sequence represents
CC an IGD-motif peptide that is specifically claimed in the invention.
XX
SQ Sequence 6 AA;
Query Match 100.0%; Score 32; DB 7; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 1 IGDIGD 6
Db 1 IGDIGD 6
RESULT 2
ADG43178
ID ADG43178 standard; peptide; 6 AA.
XX
AC ADG43178;
XX
DT 26-FEB-2004 (first entry)
XX
DE IGD-motif peptide #SEQ ID 3.
XX
KM Angiogenic composition; isoleucine-glycine-aspartic acid; IGD;
KM angiogenic growth factor; cell migration; angiogenesis; ischaemia; heart;
KM polyvinyl pyrrolidone.
XX
OS Unidentified.
XX
PN WO2003059436-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US041484.
XX
PR 21-DEC-2001; 2001US-00027015.
XX
PA (CENT-) CENTERPULSE BIOLOGICS INC.
XX
PI Akella R, Ranieri J;
XX
PT WPI; 2003-697294/66.
XX
DR
XX
PT Angiogenic composition comprising peptides containing isoleucine-glycine-
PT aspartic acid, that stimulate cell migration, and angiogenic growth
PT factor useful for promoting myocardial or peripheral angiogenesis.
XX
PS Claim 2; SEQ ID NO 3; 22pp; English.
XX
XX The invention relates to an angiogenic composition (I) comprising at
CC least one peptide chosen from a group of peptides containing isoleucine-
CC glycine-aspartic acid (IGD-containing peptide), and at least one
CC angiogenic growth factor other than the at least one peptide, where the
CC peptide stimulates cell migration. Also disclosed is a composition (III)
CC that is active for promoting cell migration and/or angiogenesis under
CC cell growth promoting conditions comprising an IGD-containing peptide,
CC and a matrix material. Compositions of the invention are useful for
CC promoting myocardial angiogenesis which involves administering (I) or
CC (III) intramyocardially to an ischaemic area of the heart of an
CC individual, to enhance vascular endothelial cell migration and/or
CC proliferation. They are also useful for promoting peripheral angiogenesis

CC in an ischaemic region of an organ or tissue fed by a peripheral vessel,
CC by enhancing vascular endothelial cell migration and/or proliferation at
CC the ischaemic region, or for enhancing blood flow to an ischaemic tissue
CC of the body, where the composition stimulates vascular endothelial cell
CC migration and/or proliferation sufficient to restore or increase blood
CC flow to the ischaemic tissue. The physiologically acceptable carrier in
CC the composition is polyvinyl pyrrolidone. The current sequence represents
CC an IGD-motif peptide that may be utilised in an angiogenic composition of
CC the invention.
XX
SQ Sequence 6 AA;
Query Match 100.0%; Score 32; DB 7; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 1 IGDIGD 6
Db 1 IGDIGD 6
RESULT 3
ABP65456
ID ABP65456 standard; protein; 443 AA.
XX
AC ABP65456;
XX
DT 19-NOV-2002 (first entry)
XX
DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:200.
XX
KM Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KM anti-diarrheic; antibacterial; inhibitor of Salmonella; detection;
KM identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KM rotavirus; food composition; pharmaceutical composition.
XX
OS Bifidobacterium longum.
XX
PN EPI227152-A1.
XX
PD 31-JUL-2002.
XX
PF 30-JAN-2001; 2001EP-00102050.
XX
PR 30-JAN-2001; 2001EP-00102050.
XX
PA (NEST) SOC PROD NESTLE SA.
XX
DR WPI; 2002-668397/72.
XX
PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as
PT a probe or primer for detecting and/or identifying Bifidobacterium longum
PT in a biological sample.
XX
PS Claim 3; SEQ ID NO 200; 80pp; English.
XX
XX The present invention describes a polynucleotide (I) comprising a
CC sequence of a Bifidobacterium genome selected from the nucleotide
CC sequences given in AB081842 and AB081843, or a sequence exhibiting at
CC least 90% identity or which hybridises with the sequence given in
CC AB081842 and AB081843. Also described is a polynucleotide (II) encoding a
CC fusion protein, comprising a sequence selected from 1097 sequences given
CC in ABP65258 to ABP65354 ligated in frame to a polynucleotide encoding a
CC heterologous polypeptide. (I) has anti-diarrheic and antibacterial
CC activities, and can be used as an inhibitor of Salmonella. (I) (which is
CC a probe) is useful for the detection and/or identification of
CC Bifidobacterium longum in a biological sample. A carrier containing the
CC lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM 1-2618) can be
CC used for preventing and/or treating diarrhoea brought about by pathogenic
CC bacteria and/or rotavirus. The carrier is a food composition selected
CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC products, ice-creams, fermented cereal based products, milk based
CC powders, infant formula, pet food or a pharmaceutical composition

CC selected from tablets, liquid bacterial suspensions, dried oral
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
 CC (1) is useful in DNA arrays or chips to carry out analysis of the
 CC expression of the *Bifidobacterium* gene. AB081844 to AB081850 represent
 CC *Bifidobacterium* related nucleotide sequences given in the Sequence
 CC Listing from the present invention but not mentioned further within the
 CC specification. N.B. The sequence data for this patent is not represented
 CC in the printed specification but is based on sequence information
 CC supplied by the European Patent Office

XX Sequence 443 AA;

Query Match 100.0%; Score 32; DB 5; Length 443;
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
 Db 155 IGDIGD 160

RESULT 4
 ADS29151 ID ADS29151 standard; protein; 515 AA.

XX ADS29151;

XX 02-DEC-2004 (first entry)

XX Bacterial polypeptide #18184.

XX Recombinant DNA construct; transformed plant; improved plant property;
 XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 XX pathogen tolerance; pest tolerance; plant disease resistance;
 XX cell cycle pathway modification; plant growth regulator;
 XX homologous recombination; seed oil yield; protein yield; carbohydrate;
 XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 XX bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
 XX for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 18184; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 XX promoter functional in a plant cell, where the promoter is positioned to
 XX provide for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source. The invention also relates to a transformed plant
 XX comprising the recombinant DNA construct and a method of producing a
 XX transformed plant having an improved property. The plant is a crop plant
 XX such as maize or soybean. The method of producing a transformed plant
 XX having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 515 AA;

Query Match 100.0%; Score 32; DB 8; Length 515;
 Best Local Similarity 100.0%; Pred. No. 6.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
 Db 423 IGDIGD 428

RESULT 5
 AAW22623 ID AAW22623 standard; protein; 59 AA.

XX AAW22623;

XX 27-AUG-2003 (revised)

XX 25-MAR-2003 (revised)

XX 12-NOV-1997 (first entry)

XX Type II topoisomerase database reference sequence SEQ ID NO.7.

XX Detection, identification, TopoII, contamination; food; human; animal;

XX plant; soil; water.

XX *Rickettsia rickettsii*.

XX Key Location/Qualifiers

XX Region 1..8

XX /note= "Variable segment 2 of type II topoisomerase"

XX Region 9..52

XX /note= "Variable segment 6 of type II topoisomerase"

XX Region 53..59

XX /note= "Variable segment 4 of type II topoisomerase"

XX US5645994-A.

XX 08-JUL-1997.

XX 06-JUN-1995; 95US-00470179.

XX 05-JUL-1990; 90US-00548138.

XX 13-AUG-1993; 93US-00106482.

XX (UTAH) UNIV UTAH RES FOUND.

XX Huang WM;

XX WPI; 1997-362925/33.

XX Detection and identification of organisms - using DNA primers to amplify
 XX signature segment of organism's type II topoisomerase.

XX Claim 8; Col 21-22; 114pp; English.

CC A method has been produced for selectively amplifying DNA segments of one
 CC or more species of organisms in a sample. The method involves: (a)
 CC providing a database containing reference sequences, comprising a subunit
 CC sequence of a signature region of a macromolecule selected from a type II
 CC topoisomerase (TopoII), or a homologue, where each reference sequence is
 CC specific to a different species of a chosen group, and the macromolecule
 CC comprises 1st and 2nd conserved regions adjacently flanking the signature
 CC region; and (b) making an extract of DNA molecules, and selectively
 CC amplifying DNA segments of the signature region using a universal primer
 CC composition, comprising a primer constructed to bind a DNA encoding the
 CC macromolecule, to produce amplified DNA segments. The present sequence
 CC represents a protein fragment containing variable portions of the type II
 CC topoisomerase for *Rickettsia rickettsii* for use in the database. The
 CC method can be used to identify all of the different organisms present in
 CC a single sample without using multiple probes. It can accurately
 CC distinguish among similar and related species, and can be used with, e.g.,
 CC contaminated food products, tissue or fluid samples from diseased humans,
 CC animal or plants, soil samples and water samples from any source.
 CC (Updated on 25-MAR-2003 to correct PF field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)
 CC
 SQ Sequence 59 AA;

Query Match 96.9%; Score 31; DB 2; Length 59;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ICDIGD 6
 |||||
 7 ICDVGD 12

RESULT 6
 AAG01407
 ID AAG01407 standard; protein; 69 AA.

AC AAG01407;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 5488.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-00200610.

PR 26-FEB-1999; 99US-0122487P.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

DR N-PSDB; AAC01413.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

PS Claim 13; SEQ ID NO 5488; 71pp + Sequence Listing; English.

CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT

CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors
 CC
 SQ Sequence 69 AA;

Query Match 96.9%; Score 31; DB 3; Length 69;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ICDIGD 6
 |||||
 40 VGDIGD 45

RESULT 7
 AB055461
 ID AB055461 standard; protein; 81 AA.

AC AB055461;

DT 29-JUL-2004 (first entry)

DE Human genome derived single exon protein #1695.

KW Human; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.

OS Homo sapiens.

PN US2003194704-A1.

PD 16-OCT-2003.

PF 03-APR-2002; 2002US-00029386.

PR 03-APR-2002; 2002US-00029386.

PA (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.

PI Penn SG, Rank DR, Hanzel DK;

PD WPI; 2004-119264/12.

PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.

PS Claim 45; SEQ ID NO 29095; 80pp; English.

CC The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid

CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing gross
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704

CC XX Sequence 81 AA;

Query Match 96.9%; Score 31; DB 8; Length 81;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 40 VGDIGD 45

RESULT 8
AA001408
ID AA001408 standard; protein; 86 AA.

XX AC AAG01408;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein, SEQ ID NO: 5489.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-00200610.

XX PR 26-FEB-1999; 99US-0122487P.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclet A, Giordano J;

XX DR MPI; 2000-500381/45.

XX N-PSDB; AAC01414.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX PS Claim 13; SEQ ID NO 5489; 71bp + Sequence Listing; English.

XX CC The present sequence is a polypeptide encoded by one of a large number of
XX 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
XX prepared from total human RNAs or polyA+ RNAs derived from 30 different
XX tissues. EST sequences usually correspond mainly to the 3' untranslated
XX region (UTR) of the mRNA because they are often obtained from oligo-dT

CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
CC sequences derived from the 5' ends of mRNAs and even in those cases where
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
CC are also used in diagnostic, forensic, gene therapy and chromosome
CC mapping procedures. They are used to obtain upstream regulatory sequences
CC and to design expression and secretion vectors

CC XX Sequence 86 AA;

Query Match 96.9%; Score 31; DB 3; Length 86;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 57 VGDIGD 62

RESULT 9
AAU18547
ID AAU18547 standard; protein; 129 AA.

XX AC AAU18547;

XX DT 21-NOV-2001 (first entry)

XX DE Human cytoskeletal element-related polypeptide #40.

XX KW Cytoskeletal element-related protein; human; mouse; rabbit; goat; horse;
XX cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
XX antirheumatic; antiproliferative; cyostatic; cardiant; neuroprotective;
XX cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
XX ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;
XX hyperproliferative disorder; breast; liver; cardiovascular disorder;
XX cerebrovascular disorder; nervous system disorder; bacterial infection;
XX fungal infection; viral infection; ocular disorder; endocrine disorder;
XX gastrointestinal disorder; renal disorder; respiratory disorder;
XX wound healing; skin aging; organ transplantation; food preservative;
XX tissue regeneration; anti-infertility; food additive.

XX OS Homo sapiens.

XX PN WO200155168-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001331.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249287P.
PR 17-NOV-2000; 2000US-0249289P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
PI WPI; 2001-476182/51.
XX N-PSDB; AAS29809.
XX
XX Novel isolated human cytoskeletal element-related polypeptide useful for
PT diagnosis/treatment of neoplastic disorders, disorders associated with
PT neural transmission, chromosomal abnormalities, autoimmune disorders.
XX
PS Claim 11; SEQ ID NO 94; 505bp; English.
XX
XX Sequences AAU18508-AAU18511 represent the cytoskeletal element-related
CC polypeptides of the invention. Cytoskeletal polypeptides and their
CC associated polynucleotides are useful in the diagnosis, treatment and
CC prevention of various types of disorders in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. A pathological condition
CC can be determined by determining the presence or absence of a mutation in
CC a cytoskeletal polynucleotide. The treatable disorders include autoimmune
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
CC as neoplasms of the breast or liver, cardiovascular disorders such as
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
CC nervous system disorders such as Alzheimer's disease, infections caused
CC by bacteria, viruses and fungi, ocular disorders such as corneal
CC infection, endocrine disorders such as premature labour and infertility,
CC gastrointestinal disorders such as Crohn's disease, renal disorders such
CC as glomerulonephritis and respiratory disorders such as asthma. The
CC polypeptides can also be used to aid wound healing, to prevent skin aging
CC due to sunburn, to maintain organs before transplantation, to regenerate
CC tissues and in chemotaxis. The polypeptides can also be used as a food
CC additive or preservative to increase or decrease storage capabilities.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 96.9%; Score 31; DB 4; Length 129;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 69 VGDIGD 74

RESULT 10

AAU51226
ID AAU51226 standard; protein; 149 AA.

AAU51226;

27-FEB-2002 (first entry)

Propionibacterium acnes immunogenic protein #12122.

SAPHO syndrome; synovitis; acne; pustulosis; hypertosia; osteomyelitis;
uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
dermatological; osteopathic; neuroprotectant.

Propionibacterium acnes.

WO200181581-A2.

01-NOV-2001.

20-APR-2001; 2001WO-US012865.

21-APR-2000; 2000US-0199047P.

02-UN-2000; 2000US-0208841P.

07-JUL-2000; 2000US-0216747P.

(CORI-) CORIXA CORP.

Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

L'maisonneuve J, Zhang Y, Jen S, Carter D;

WPI; 2001-616774/71.

N-PSDB; AAS59551.

Propionibacterium acnes polypeptides and nucleic acids useful for
vaccinating against and diagnosing infections, especially useful for
treating acne vulgaris.

Example 1; SEQ ID NO 12421; 1069pp; English.

Sequences AAU3105-AAU68017 represent Propionibacterium acnes immunogenic
polypeptides. The proteins and their associated DNA sequences are used in
the treatment, prevention and diagnosis of medical conditions caused by
P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
pustulosis, hypertosia and osteomyelitis), uveitis and endophthalmitis.
P. acnes is also involved in infections of bone, joints and the central
nervous system, however it is particularly involved in the inflammatory
lesions associated with acne vulgaris. A method for detecting the
presence or absence of P. acnes in a patient comprises contacting a
sample with a binding agent that binds to the proteins of the invention
and determining the amount of bound protein in the sample. The
polypeptides may be used as antigens in the production of antibodies
specific for P. acnes proteins. These antibodies can be used to
downregulate expression and activity of P. acnes polypeptides and
therefore treat P. acnes infections. The antibodies may also be used as
diagnostic agents for determining P. acnes presence, for example, by
enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
this patent did not form part of the printed specification, but was
obtained in electronic format directly from WIPO at

ftp.wipo.int/pub/published_pct_sequences

Query Match 96.9%; Score 31; DB 4; Length 149;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 49 VGDIGD 54

RESULT 11

ABM47745
ID ABM47745 standard; protein; 149 AA.

ABM47745;

20-OCT-2003 (first entry)

Propionibacterium acnes predicted ORF-encoded polypeptide #12421.

Acne vulgaris; antiseborrheic; dermatological; antibacterial;
immunostimulant; immune response; vaccine.

Propionibacterium acnes.

WO2003033515-A1.

24-APR-2003.

11-OCT-2002; 2002WO-US032727.

15-OCT-2001; 2001US-00978825.

(CORI-) CORIXA CORP.

Mitcham JL, Skelky YAW, Persing DH, Bhatia A, Maisonneuve JL;

Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

Barth B, Vallieve-Douglas J;

WPI; 2003-381789/36.

N-PSDB; ACP64480.

New Propionibacterium acnes polypeptides and polynucleotides encoding the
polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
or for stimulating an immune response specific for a P. acnes protein.

Example 1; SEQ ID NO 12421; 1481pp; English.

The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of a P. acnes
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present

CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPD at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 149 AA;

Query Match 96.9%; Score 31; DB 6; Length 149;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
: |||||
Db 49 VGDIGD 54

RESULT 12
AAU18521
ID AAU18521 standard; protein; 156 AA.

XX AAU18521;

DT 21-NOV-2001 (first entry)

XX Human cytoskeletal element-related polypeptide #14.

XX Cytoskeletal element-related protein; human; mouse; rabbit; goat; horse;
XX cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
XX antirheumatic; antiproliferative; cytostatic; cardiatic; neuroprotective;
XX cerebroprotective; neurotropic; antibacterial; virocidic; fungicide; cancer;
XX ophthalmological; vulnerrary; gene therapy; autoimmune disease; neoplasm;
XX hyperproliferative disorder; breast; liver; cardiovascular disorder;
XX cerebrovascular disorder; nervous system disorder; bacterial infection;
XX fungal infection; viral infection; ocular disorder; endocrine disorder;
XX gastrointestinal disorder; renal disorder; respiratory disorder;
XX wound healing; skin aging; organ transplantation; food preservative;
XX tissue regeneration; anti-infertility; food additive.

OS Homo sapiens.

XX MO200155168-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001MO-US001331.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216860P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM,
XX
XX WPI: 2001-476182/51.
DR N-PSDB; AAS29783.
XX
XX
PT Novel isolated human cytoskeletal element-related polypeptide useful for
PI diagnosis/treatment of neoplastic disorders, disorders associated with
PI neural transmission, chromosomal abnormalities, autoimmune disorders.
XX
XX
PS Claim 11; SEQ ID NO 68; 505pp; English.
XX
XX Sequences AAU98508-AAU16551 represent the cytoskeletal element-related
XX polypeptides of the invention. Cytoskeletal polypeptides and their
XX associated polynucleotides are useful in the diagnosis, treatment and
XX prevention of various types of disorders in e.g. humans, mice, rabbits,
XX goats, horses, cats, dogs, chickens or sheep. A pathological condition
XX can be determined by determining the presence or absence of a mutation in
XX a cytoskeletal polynucleotide. The treatable disorders include autoimmune
XX diseases such as rheumatoid arthritis, hyperproliferative disorders such
XX as neoplasms of the breast or liver, cardiovascular disorders such as
XX cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
XX nervous system disorders such as Alzheimer's disease, infections caused
XX by bacteria, viruses and fungi, ocular disorders such as corneal
XX infection, endocrine disorders such as premature labour and infertility,
XX gastrointestinal disorders such as Crohn's disease, renal disorders such
XX as glomerulonephritis and respiratory disorders such as asthma. The
XX polypeptides can also be used to aid wound healing, to prevent skin aging
XX due to sunburn, to maintain organs before transplantation, to regenerate
XX tissues and in chemotaxis. The polypeptides can also be used as a food
XX additive or preservative to increase or decrease storage capabilities.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences

Query Match 96.9%; Score 31; DB 4; Length 166;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 IGDIGD 6
Db 101 VGDIGD 106
RESULT 13
AAM78545
ID AAM78545 standard; protein, 166 AA.
XX
XX AAM78545;
XX
XX
XX 06-NOV-2001 (first entry)
XX
XX Human protein SEQ ID NO 1207.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation.
XX
XX Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001MO-US004098.
XX
XX 03-FEB-2000; 2000US-00496914.
XX 27-APR-2000; 2000US-00560875.
XX 20-JUN-2000; 2000US-00598075.
XX 19-JUL-2000; 2000US-00620325.
XX 01-SEP-2000; 2000US-00654936.
XX 15-SEP-2000; 2000US-00653551.
XX 20-OCT-2000; 2000US-00693345.
XX 30-NOV-2000; 2000US-00728422.
XX
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
XX Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZM;
XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR N-PSDB; AAK51678.
XX
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX in diagnosis and gene therapy.
XX
XX
XX Claim 20; Page 3461-3462; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication
XX
XX Sequence 166 AA;
XX
XX Query Match 96.9%; Score 31; DB 4; Length 166;

Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
:|||||
Db 57 VGDIGD 62

RESULT 14

ABR41956
ID ABR41956 standard; protein; 166 AA.

AC ABR41956;

DT 11-AUG-2003 (first entry)

DE Mouse cofilin, implicated in pain.

KW Mouse; cofilin; nociceptive; analgesic; signal transduction;
transgenic animal.

OS Mus musculus.

PN EP1281775-A2.

PD 05-FEB-2003.

PF 26-JUL-2002; 2002EP-00255231.

PR 27-JUL-2001; 2001GH-00018354.

PR 07-FEB-2002; 2002GB-00002892.

PA (WARN) WARNER LAMBERT CO.

PI Brookbank RA, Dixon AK, Lee K, Plimock RD;

DR WPI; 2003-335009/32.

DR N-PSDB; ACC48821.

PT Use of isolated gene or nucleic acid sequence, recombinant vector, host
cell, non-human animal, polypeptide encoded by the nucleic acid sequence,
PT or antibody, for screening of compounds for the treatment of pain, or for
diagnosing pain.

PS Disclosure; Page 65-66; 87pp; English.

CC The present sequence is the protein sequence of mouse cofilin. This is
CC encoded by a gene that has been identified as being up-regulated in 2
CC models of chronic pain, i.e. streptozocin-induced diabetes and chronic
CC constitutive injury to a nerve leading to the spine. The expression
CC products of such genes can be used to screen libraries for compounds and
CC peptide agonists and antagonists of gene product activity that may be
CC useful in the treatment or prevention of chronic pain, and in the
CC development of diagnostic tools for the identification and
CC characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

CC Characterisation of pain

DT 04-DEC-2003 (first entry)

DE Prostate cancer marker protein.

KW Prostate; cancer; cytostatic; gene therapy; marker.

OS Homo sapiens.

PN WO2003009814-A2.

PD 06-FEB-2003.

PF 25-JUL-2002; 2002MO-US023913.

PR 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 03-MAR-2002; 2002US-0362158P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersch S, Kamatkar S, Monsey AM, Glatc K, Zhao X, Anderson D;

DR WPI; 2003-248033/24.

PT New nucleic acid molecule, useful for diagnosing or treating prostate
cancer.

PS Disclosure; SEQ ID NO 72; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.

CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

DT 04-DEC-2003 (first entry)

DE Prostate cancer marker protein.

KW Prostate; cancer; cytostatic; gene therapy; marker.

OS Homo sapiens.

PN WO2003009814-A2.

PD 06-FEB-2003.

PF 25-JUL-2002; 2002MO-US023913.

PR 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 03-MAR-2002; 2002US-0362158P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersch S, Kamatkar S, Monsey AM, Glatc K, Zhao X, Anderson D;

DR WPI; 2003-248033/24.

PT New nucleic acid molecule, useful for diagnosing or treating prostate
cancer.

PS Disclosure; SEQ ID NO 72; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.

CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

DT 04-DEC-2003 (first entry)

DE Prostate cancer marker protein.

KW Prostate; cancer; cytostatic; gene therapy; marker.

OS Homo sapiens.

PN WO2003009814-A2.

PD 06-FEB-2003.

PF 25-JUL-2002; 2002MO-US023913.

PR 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 03-MAR-2002; 2002US-0362158P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersch S, Kamatkar S, Monsey AM, Glatc K, Zhao X, Anderson D;

DR WPI; 2003-248033/24.

PT New nucleic acid molecule, useful for diagnosing or treating prostate
cancer.

PS Disclosure; SEQ ID NO 72; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.

CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

DT 04-DEC-2003 (first entry)

DE Prostate cancer marker protein.

KW Prostate; cancer; cytostatic; gene therapy; marker.

OS Homo sapiens.

PN WO2003009814-A2.

PD 06-FEB-2003.

PF 25-JUL-2002; 2002MO-US023913.

PR 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 03-MAR-2002; 2002US-0362158P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersch S, Kamatkar S, Monsey AM, Glatc K, Zhao X, Anderson D;

DR WPI; 2003-248033/24.

PT New nucleic acid molecule, useful for diagnosing or treating prostate
cancer.

PS Disclosure; SEQ ID NO 72; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.

CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

DT 04-DEC-2003 (first entry)

DE Prostate cancer marker protein.

KW Prostate; cancer; cytostatic; gene therapy; marker.

OS Homo sapiens.

PN WO2003009814-A2.

PD 06-FEB-2003.

PF 25-JUL-2002; 2002MO-US023913.

PR 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 03-MAR-2002; 2002US-0362158P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersch S, Kamatkar S, Monsey AM, Glatc K, Zhao X, Anderson D;

DR WPI; 2003-248033/24.

PT New nucleic acid molecule, useful for diagnosing or treating prostate
cancer.

PS Disclosure; SEQ ID NO 72; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.

CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

CC Sequence 166 AA;

DT 04-DEC-2003 (first entry)

DE Prostate cancer marker protein.

KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cyostatic.
 XX Homo sapiens.
 OS
 PN WO2003087768-A2.
 XX
 PD 23-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010870.
 XX
 PR 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-0399987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX
 PA (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.
 PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Watnock DE;
 PI
 DR WPI; 2003-845369/78.
 XX
 XX Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function;
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 XX
 XX Claim 1; SEQ ID NO 1910; 180pp; English.
 XX
 CC This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nootropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cyostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.
 XX
 SQ Sequence 166 AA;
 XX
 Query Match 96.9%; Score 31; DB 7; Length 166;
 Best Local Similarity 83.3%; Pred. No. 3.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IGDIGD 6
 :|||||
 DB 57 VGDIGD 62
 XX
 RESULT 17
 ADI24536
 ID ADI24536 standard; protein; 166 AA.
 XX
 AC ADI24536;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Human modifier of Chk1 (MCHK) protein SEQ ID NO:86.
 XX
 KW Chk1 pathway modulating agent; modifier of Chk1; MCHK; cyostatic;
 KW gene therapy; cancer; human.
 XX
 OS Homo sapiens.
 XX

PN WO2004004785-A1.
 XX
 PD 15-JAN-2004.
 XX
 PF 09-JUL-2003; 2003WO-US021379.
 XX
 PR 10-JUL-2002; 2002US-0394845P.
 PR 16-SEP-2002; 2002US-0410986P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 PI Francis-Lang H, Roche S, Joo DM, Nicoll M, Hai B, Zhang H;
 PI Lickteig K, Amundsen CD, Jin Y, Adamkewicz JI, Platt DM;
 PI Hammonds RG;
 PI
 DR WPI; 2004-083465/08.
 DR N-PSDB; ADI24486.
 XX
 XX Identifying a candidate Chk1 pathway modulating agent for treating e.g.,
 PT cancer, comprises contacting an assay system comprising a MCHK
 PT polypeptide or nucleic acid with a test agent and detecting a test agent-
 PT biased activity.
 XX
 XX Example; SEQ ID NO 86; 266pp; English.
 PS
 CC The present invention describes a method for identifying a candidate Chk1
 CC pathway modulating agent. The method comprises: (a) providing an assay
 CC system comprising a modifier of Chk1 (MCHK) polypeptide or nucleic acid;
 CC (b) contacting the system with a test agent, where the system provides a
 CC reference activity except in the presence of the test agent; and (c)
 CC detecting a test agent-biased activity, and a difference between the test
 CC agent-biased activity and the reference activity. Also described: (1) a
 CC method for modulating Chk1 pathway of a cell; (2) a method for modulating
 CC Chk1 pathway in a mammalian cell; and (3) a method for diagnosing a
 CC disease in a patient. A MCHK sequence has cyostatic activity, and can be
 CC used in gene therapy. The method is useful for identifying a candidate
 CC Chk1 pathway-modulating agent for preparing a composition for diagnosing
 CC or treating e.g., cancer. The present sequence represents a human MCHK
 CC protein, which is used in the exemplification of the present invention.
 XX
 SQ Sequence 166 AA;
 XX
 Query Match 96.9%; Score 31; DB 8; Length 166;
 Best Local Similarity 83.3%; Pred. No. 3.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IGDIGD 6
 :|||||
 DB 57 VGDIGD 62
 XX
 RESULT 18
 AAU52312
 ID AAU52312 standard; protein; 173 AA.
 XX
 AC AAU52312;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #13208.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX

XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'atsomeuve J, Zhang Y, Jen S, Carter D;
PI N-PSDB; AAS59554.
DR WPI; 2001-616774/71.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 13507; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPPO syndrome (synovitis, acne,
CC pustulosis, hysterosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 173 AA;
XX
Query Match 96.9%; Score 31; DB 4; Length 173;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 IGDIGD 6
Db 102 IGDVGD 107
XX
RESULT 19
ABM48831
ID ABM48831 standard; protein; 173 AA.
XX
AC ABM48831;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #13507.
XX
KW Acne vulgaris; antisephorhoic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
XX
PN W02003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002MO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Malsomeuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valiieve-Douglas J;
XX
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64483.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 13507; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 173 AA;
XX
Query Match 96.9%; Score 31; DB 6; Length 173;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 IGDIGD 6
Db 102 IGDVGD 107
XX
RESULT 20
ADQ10039
ID ADQ10039 standard; protein; 201 AA.
XX
AC ADQ10039;
XX
DT 07-OCT-2004 (first entry)
XX
DE B_afzelii ORF 13 protein product of the vls locus silent cassette Seq 83.
XX
KW variable major protein; VMP; VMP-like sequence; VLS; Lyme disease;
KW relapsing fever; antibacterial; gene therapy; immunoprophylaxis;
KW VMP antigenic variation system; pathogenic.
XX
OS Borrelia afzelii.
XX
XX
PN W02004058181-A2.
XX

PR 29-SEP-2000; 2000US-0236359P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-476182/51.
XX N-PSDB; AAS29808.
DR

XX Novel isolated human cytoskeletal element-related polypeptide useful for
PT diagnosis/treatment of neoplastic disorders, disorders associated with
PT neural transmission, chromosomal abnormalities, autoimmune disorders.
XX
PS Claim 11; SEQ ID NO 93; 505pp; English.
XX
CC Sequences AAU18508-AAU18551 represent the cytoskeletal element-related
CC polypeptides of the invention. Cytoskeletal polypeptides and their
CC associated polynucleotides are useful in the diagnosis, treatment and
CC prevention of various types of disorders in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. A pathological condition
CC can be determined by determining the presence or absence of a mutation in
CC a cytoskeletal polynucleotide. The creatable disorders include autoimmune
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
CC as neoplasms of the breast or liver, cardiovascular disorders such as
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
CC nervous system disorders such as Alzheimer's disease, infections caused
CC by bacteria, viruses and fungi, ocular disorders such as corneal
CC infection, endocrine disorders such as premature labour and infertility,
CC gastrointestinal disorders such as Crohn's disease, renal disorders such
CC as glomerulonephritis and respiratory disorders such as asthma. The
CC polypeptides can also be used to aid wound healing, to prevent skin aging
CC due to sunburn, to maintain organs before transplantation, to regenerate
CC tissues and in chemotaxis. The polypeptides can also be used as a food
CC additive or preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)
Query Match 96.9%; Score 31; DB 4; Length 205;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 IGDIGD 6
Db 96 VGDIGD 101
RESULT 22
AAW79529
ID AAW79529 standard; protein; 207 AA.
XX
AC AAW79529;
XX
DT 06-NOV-2001 (first entry)
XX
XX Human protein SEQ ID NO 3175.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
XX WO200157190-A2.
XX
PD 09-AUG-2001.
XX
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
XX
XX 03-FEB-2000; 2000US-00496914.
XX 27-APR-2000; 2000US-00560875.
XX 20-JUN-2000; 2000US-00598075.
XX 19-JUL-2000; 2000US-00620325.
XX 01-SEP-2000; 2000US-00654936.
XX 15-SEP-2000; 2000US-00663561.
XX 20-OCT-2000; 2000US-00693325.
XX 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX

PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y,
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW,
PI Xue HJ, Yang Y, Wejthman T, Goodrich R;
XX WPI; 2001-476283/51.
DR N-PSDB; AAK52662.
XX
PT Nucleic acid encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 275; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 207 AA;
XX
Query Match 96.9%; Score 31; DB 4; Length 207;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 IGDIGD 6
:|||||
DB 98 VGDIGD 103
XX
RESULT 23
ADQ10053
ID ADQ10053 standard; protein; 220 AA.
XX
AC ADQ10053;
XX
DT 07-OCT-2004 (first entry)
XX
DE B_afzeili strain ACAI vls 13 protein expressed by plasmid pBA-13-1 Seq97.
XX
KM variable major protein; VMP; VMP-like sequence; VLS; Lyme disease;
KM relapsing fever; antibacterial; gene therapy; immunophylaxis;
XX VMP antigenic variation system; pathogenic.
OS Borrelia afzeili.
XX
PN WO2004058181-A2.
XX
PD 15-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041182.
XX
PR 20-DEC-2002; 2002US-0435077P.
XX
PI (TEXA) UNITV TEXAS SYSTEM.
XX
PA Norris SJ;
XX
PI WPI; 2004-525783/50.
DR N-PSDB; ADQ10052.
XX
DR New isolated nucleic acid molecules encoding Variable Major Protein-like
PT polypeptides of Borrelia, useful for diagnosing, preventing or treating
PT Borrelia infections such as Lyme disease or relapsing fever.
XX
PS Claim 22; SEQ ID NO 97; 182pp; English.

XX
CC This invention relates to a novel nucleic acid molecules isolated from
CC pathogenic Borrelia bacteria that each encode a variable major protein
CC (VMP)-like sequence (VLS) peptide. Specifically, it refers to recombinant
CC VLS genes and immunogenic compositions derived thereof, from Borrelia
CC garinii (B. garinii) and Borrelia afzeili (B. afzeili). The present
CC invention describes the use of antibodies against the expressed VLS
CC proteins where immunologic binding is indicative of a Borrelia infection,
CC in particular infections that include Lyme disease, relapsing fever or
CC related diseases in humans and animals. Accordingly, these antibacterial
CC compositions can be used for gene therapy, immunophylaxis, diagnosis
CC or treatment of such diseases, which due to the VMP antigenic variation
CC system often evade the immune system response despite the occurrence of
CC host antibody and cellular responses. This polypeptide is the Borrelia
CC afzeili strain ACAI vls ORF 13 protein expressed by the pBA-13-1 plasmid
CC of the invention.
XX
SQ Sequence 220 AA;
XX
Query Match 96.9%; Score 31; DB 8; Length 220;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 IGDIGD 6
:|||||
DB 46 IGDVGD 51
XX
RESULT 24
ADJ48363
ID ADJ48363 standard; protein; 342 AA.
XX
AC ADJ48363;
XX
DT 06-MAY-2004 (first entry)
XX
DE Maize oil-associated gene protein #22.
XX
KM maize; plant; oil-associated gene; transgenic; enhanced seed oil;
KM vegetable oil.
XX
OS Zea mays.
XX
PN US2004025202-A1.
XX
PD 05-FEB-2004.
XX
PF 14-MAR-2003; 2003US-00389566.
XX
PR 15-MAR-2002; 2002US-0365301P.
PR 26-JUN-2002; 2002US-0391786P.
PR 26-JUN-2002; 2002US-0392018P.
XX
PA (LAUR/) LAURIE C C.
PA (RAVA/) RAVANELLO M.
PA (SAVA/) SAVAGE T.
PA (LEDE/) LEDEAUX J R.
PA (ROGE/) ROGERS J A.
XX
PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
XX
PI WPI; 2004-142683/14.
XX
PT Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.
XX
XX Disclosure; SEQ ID NO 367; 22pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in plants operably linked to an oil-associated gene.
CC The construct is useful for transgenic plant seed which has in its genome
CC the construct, that is functional in the plant to transcribe the oil-

CC associated gene. The transgenic plant seed grows into a plant having
CC enhanced seed oil as compared to wild type. The construct is useful for
CC producing hybrid maize seed. The transgenic plant seed is useful for
CC producing vegetable oil. The present sequence represents the amino acid
CC sequence of a maize oil-associated gene protein.
XX
SQ Sequence 342 AA;
Query Match 96.3%; Score 31; DB 8; Length 342;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIID 6
Db 51 VGDIGD 56
RESULT 25
ADS45014
ID ADS45014 standard; protein; 364 AA.
XX
AC ADS45014;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #23444.
XX
KM Recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KM pathogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant growth regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KM bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 23444; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 364 AA;
Query Match 96.3%; Score 31; DB 8; Length 364;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIID 6
Db 70 VGDIGD 75
RESULT 26
ADS22843
ID ADS22843 standard; protein; 693 AA.
XX
AC ADS22843;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #11876.
XX
KM Recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KM pathogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant growth regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KM bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 11876; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a

transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomanan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

CC Sequence 693 AA;

Query Match 96.9%; Score 31; DB 8; Length 693;

Best Local Similarity 83.3%; Pred. No. 1.4e+03; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Indels 0; Gaps 0;

OY 1 IGDIGD 6
:|||||
Db 409 IGDVGD 414

RESULT 27

AAB22942 AAB22942 standard; protein; 1971 AA.

XX AAB22942;

XX 10-JAN-2001 (first entry)

XX Mouse GANP protein.

XX GANP; mouse; murine; kinase activity; protein phosphorylation;

XX signal transduction; abnormal B cell differentiation; antibody;

XX autoimmune disorder.

XX Mus sp.

XX WO200050611-A1.

XX 31-AUG-2000.

XX 27-AUG-1999; 99WO-JP004634.

XX 24-FEB-1999; 99JP-00047035.

XX (SUME) SUMITOMO ELECTRIC IND CO.

XX Sakaguchi N, Kuwahara K;

XX WPI: 2000-549411/50.

XX N-PSDB; AAA90435.

XX GANP proteins participating in signal conversion of abnormal B cell

XX differentiation in autoimmune state and having kinase activity, useful in

XX the study of autoimmune mechanisms.

XX Claim 1; Page 58-67; 91pp; Japanese.

XX This sequence represents mouse GANP protein. The invention relates to the

XX mouse and human GANP proteins (AAB22942, AAB22943) which are 210 kD in

XX size, have kinase activity and participate in signal transduction during

XX abnormal B cell differentiation. The invention also encompasses

XX nucleotides encoding GANP proteins, variant GANP proteins, and antibodies

CC against GANP proteins. GANP proteins may be used in the study of

CC mechanisms involved in autoimmune disorders

XX Sequence 1971 AA;

XX SQ

Query Match 96.9%; Score 31; DB 3; Length 1971;

Best Local Similarity 83.3%; Pred. No. 4.2e+03; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Indels 0; Gaps 0;

OY 1 IGDIGD 6
:|||||

Db 1388 VGDIGD 1393

RESULT 28

AAB98974 AAB98974 standard; protein; 1971 AA.

XX AAB98974;

XX 20-AUG-2001 (first entry)

XX Murine SHD1 related protein.

XX Mouse; SHD1; differentiation; germ centre B cell.

XX Mus sp.

XX JP2001078778-A.

XX 27-MAR-2001.

XX 17-SEP-1999; 99JP-00263688.

XX 17-SEP-1999; 99JP-00263688.

XX (SUME) SUMITOMO ELECTRIC IND CO.

XX WPI: 2001-313371/33.

XX N-PSDB; AAH25842.

XX SHD1 protein useful for controlling the differentiation and activation of

XX germ center B cells.

XX Claim 2; Page 8-13; 16pp; Japanese.

XX The present invention relates to the murine SHD1 protein shown in

XX CC AAB98973. This protein can be used for controlling differentiation and

XX CC activation of germ centre B cells. The present sequence is a protein used

XX in the exemplification of the invention

XX SQ

Query Match 96.9%; Score 31; DB 4; Length 1971;

Best Local Similarity 83.3%; Pred. No. 4.2e+03; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Indels 0; Gaps 0;

OY 1 IGDIGD 6
:|||||

Db 1388 VGDIGD 1393

RESULT 29

AAB98957 AAB98957 standard; protein; 1971 AA.

XX AAB98957;

XX 20-AUG-2001 (first entry)

XX Murine CALP related protein.

XX Mouse; CALP; autoimmune disease.

XX Mus sp.
OS JP2001078779-A.
XX
XX 27-MAR-2001.
XX
XX 17-SEP-1999; 99JP-00263707.
XX
XX 17-SEP-1999; 99JP-00263707.
XX
XX (SUME) SUMITOMO ELECTRIC IND CO.
XX WPI; 2001-313372/33.
XX N-PSDB; AAS25794.
XX
XX New CALPP protein the use of which is related to the treatment or the
PT detection of autoimmune diseases.
XX
XX Claim 2; Page 9-13; 17pp; Japanese.
XX
XX The present invention provides a murine protein, designated CALPP and
CC shown in AAB98956. This can be used to generate an antibodies which can
CC be used for the treatment or detection of autoimmune diseases. The
CC present sequence is a protein used in the exemplification of the
CC invention
XX
XX Sequence 1971 AA;
SQ

Query Match 96.9%; Score 31; DB 4; Length 1971;
Best Local Similarity 83.3%; Pred. No. 4.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IGDIGD 6
:|||||
Db 1388 VGDIGD 1393

RESULT 30
ADO31144
ID ADO31144 standard; protein: 1971 AA.
XX
XX ADO31144;
XX
XX 12-AUG-2004 (first entry)
XX
XX Mouse germinal center associated nuclear protein.
XX
XX virucide; viral antigen inhibitor; transgenic;
XX germinal center associated nuclear protein; GAMP; antibody;
XX virus infection; human immunodeficiency virus; hepatitis C virus.
XX
XX Mus sp.
XX
XX WO2004040971-A1.
XX
XX 21-MAY-2004.
XX
XX 07-NOV-2003; 2003WO-JP014221.
XX
XX 07-NOV-2002; 2002WO-JP011598.
XX
XX (IMMU-) IMMUNOKICK INC.
XX
XX Sakaguchi N;
XX
XX WPI; 2004-411378/38.
XX N-PSDB; ADO31143.
XX
XX Transgenic mammal transformed with germinal center associated nuclear
PT protein (GAMP) gene for production of high-affinity antibodies as
PT diagnostic reagents and disease therapy.
XX

PS Disclosure, SEQ ID NO 2; 214pp; Japanese.
XX
XX The invention relates to transgenic non-human animals and their offspring
CC which are transformed with germinal center associated nuclear protein
CC (GAMP) gene. The GAMP gene, encoded protein and transgenic animals
CC express GAMP can be used for the production of high-affinity antibodies
CC to viral antigens for treatment and prevention of infection by viruses
CC such as human immunodeficiency virus and hepatitis C virus. This sequence
XX corresponds to the mouse GAMP protein.
XX
XX Sequence 1971 AA;
SQ

Query Match 96.9%; Score 31; DB 8; Length 1971;
Best Local Similarity 83.3%; Pred. No. 4.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IGDIGD 6
:|||||
Db 1388 VGDIGD 1393

RESULT 31
AAU65282
ID AAU65282 standard; protein: 58 AA.
XX
XX AAU65282;
XX
XX 27-FEB-2002 (first entry)
XX
XX

DE Propionibacterium acnes immunogenic protein #26178.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX

OS Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX 02-JUN-2000; 2000US-0208841P.
XX 07-JUL-2000; 2000US-0216747P.
XX

(CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59663.
XX

PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX

PS Example 1; SEQ ID NO 26477; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

CC Sequence 58 AA;

Query Match 93.8%; Score 30; DB 4; Length 58;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
 Db 5 LGDIGD 10

RESULT 32

ABM61801
 ID ABM61801 standard; protein; 58 AA.

XX ABM61801;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #26477.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;

XX immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skelky YAW, Persing DH, Bhatia A, Maisonneuve JL;

XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

XX Barin B, Vallieve-Douglas J;

XX WPI; 2003-381789/36.

XX N-PSDB; ACP64592.

XX Example 1; SEQ ID NO 26477; 1481bp; English.

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a

CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for the
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

CC Sequence 58 AA;

Query Match 93.8%; Score 30; DB 6; Length 58;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
 Db 5 LGDIGD 10

RESULT 33

AAU59294
 ID AAU59294 standard; protein; 78 AA.

XX AAU59294;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #20190.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59601.

XX Example 1; SEQ ID NO 20489; 1069bp; English.

CC Sequences AAU59105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of *P. acnes* in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for *P. acnes* proteins. These antibodies can be used to
CC downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 78 AA;
Query Match 93.8%; Score 30; DB 4; Length 78;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
:|:|:|
Db 71 VGDVGD 76
RESULT 34
ABM55813
ID ABM55813 standard; protein; 78 AA.
XX
AC ABM55813;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #20489.
XX
KM Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN W02003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valiieve-Douglase J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64530.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a *P. acnes* protein.
XX
PS Example 1; SEQ ID NO 20489; 1481bp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of *P. acnes* polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a *P. acnes*
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising *P. acnes* polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of *P. acnes* in a
CC patient; and a method for inhibiting the development of *P. acnes* in a
CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a *P. acnes*
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against *P. acnes*, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the *P. acnes* polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 78 AA;
Query Match 93.8%; Score 30; DB 6; Length 78;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
:|:|:|
Db 71 VGDVGD 76
RESULT 35
ADM27005
ID ADM27005 standard; protein; 87 AA.
XX
AC ADM27005;
XX
DT 20-MAY-2004 (first entry)
XX
DE Hyperthermophile Methanopyrus kandleri protein #1611.
XX
KM Hyperthermophile; protein stability enhancement;
KW protein activity enhancement.
XX
OS Methanopyrus kandleri.
XX
PN W02003076575-A2.
XX
PD 18-SEP-2003.
XX
PF 04-MAR-2003; 2003WO-US006664.
XX
PR 04-MAR-2002; 2002US-0361742P.
PR 14-MAY-2002; 2002US-0380423P.
PR 16-SEP-2002; 2002US-0410974P.
XX
PA (FIDE-) FIDELITY SYSTEMS INC.
PA (MALY/) MALYKH A.
XX
PI Slesarev AI, Pavlov A, Pavlova N, Koz'yavkin S;
XX
DR WPI; 2003-748383/70.
DR N-PSDB; ADM27081.
XX
PT New isolated nucleic acids encoding any of about 1700 Methanopyrus
PT kandleri proteins, and the encoded proteins, useful as a medicaments or
PT as diagnostic agents.
XX
PS Claim 31; SEQ ID NO 1611; 1023bp; English.
XX
CC The invention comprises the amino acid sequence of proteins from the
CC hyperthermophile Methanopyrus kandleri. The invention also comprises the
CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
CC proteins of the invention are useful for enhancing the stability and/or
CC activity of other proteins. The Methanopyrus kandleri genome is useful in

CC a variety of diagnostic and analytical methods. The present amino acid
CC sequence represents a Methanopyrus kandleri protein of the invention.
XX
SQ Sequence 87 AA;
Query Match 93.8%; Score 30; DB 7; Length 87;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
Db 8 LGDIGD 13
RESULT 36
ADE43918
ID ADE43918 standard; protein; 122 AA.
XX
AC ADE43918;
XX
DT 26-FEB-2004 (first entry)
XX
DE Chicken WE474 protein SEQ ID NO:28.
XX
KM DP119; DP444; DP810; DP685; WE474; DP160; RA977; RA770; anti-diabetic;
KM anorectic; neuroprotective; gene therapy; pancreatic disorder; diabetes;
KM neurodegenerative disorder; hyperglycaemia; impaired glucose tolerance;
XX obesity.
XX OS Gallus gallus.
XX
PN WO200309318-A2.
XX
PD 04-DEC-2003.
XX
PF 30-MAY-2003; 2003WO-EP005700.
XX
PR 29-MAY-2002; 2002EP-00011963.
PR 17-SEP-2002; 2002EP-00020829.
XX
XX
PA (DEVE-) DEVELOPENTWICKLUNGSMIOLOGISCHE FORSCH.
XX
PI Dohmann C, Austen M;
XX
DR WPI; 2004-035045/03.
DR N-PSDB; ADE43917.
XX
PT Use of DP119, DP444, DP810, DP685, WE474, DP160, RA977 or RA770 nucleic
PT acid molecules or polypeptides, or their effectors or modulators, for
PT manufacturing an agent for the diagnosis, prevention or treatment of
PT pancreatic disorders.
XX
PS Claim 3; SEQ ID NO 28; 125pp; English.
XX
XX The present invention describes a nucleic acid molecule (I) selected from
CC DP119, DP444, DP810, DP685, WE474, DP160, RA977 and RA770, or its encoded
CC polypeptide, or a fragment or variant of the nucleic acid or polypeptide,
CC or an effector/modulator of the nucleic acid or polypeptide. (I) is used
CC in manufacturing a pharmaceutical agent, in monitoring or controlling the
CC function of a gene or gene product influenced or modified by a DP119,
CC DP444, DP810, DP685, WE474, DP160, RA977 or RA770 polypeptide, or in
CC identifying substances that interact with the polypeptide. (I) has
CC anti-diabetic, anorectic and neuroprotective activities, and can be used
CC in gene therapy. (I) is useful for diagnostic or therapeutic
CC applications, or for manufacturing an agent for diagnosis, monitoring,
CC prevention or treatment of pancreatic disorders, such as diabetes and
CC related disorders, as well as neurodegenerative disorders, and other
CC diseases. It may also be used for detecting and/or verifying, for the
CC treatment, alleviation and/or prevention, of a pancreatic dysfunction
CC (e.g. diabetes, hyperglycaemia or impaired glucose tolerance), and
CC related disorders including obesity, and neurodegenerative disorders, and
CC others, in cells, cell masses, organs and/or subjects. It is also used
CC for regenerating or for promoting the differentiation and/or function of

CC beta-cells in vitro and/or in vivo. (I) may also be used in preparing a
CC non-human animal which over- or under-expresses the DP119, DP444, DP810,
CC DP685, WE474, DP160, RA977 or RA770 gene product. In addition, the
CC nucleic acid molecule, polypeptide or an effector/modulator of the
CC nucleic acid or polypeptide, is used in monitoring or controlling the
CC function of a gene or gene product influenced or modified by a DP119,
CC DP444, DP810, DP685, WE474, DP160, RA977 or RA770 polypeptide, or in
CC identifying substances that interact with the polypeptide. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 122 AA;
Query Match 93.8%; Score 30; DB 8; Length 122;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
Db 79 VGDVGD 84
RESULT 37
AAU63370
ID AAU63370 standard; protein; 127 AA.
XX
AC AAU63370;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #24266.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
PN
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0198047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Peking DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59633.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 24565; 1069pp; English.
XX
XX Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The

CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of *P. acnes* in a
CC patient; and a method for inhibiting the development of *P. acnes* in a
CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating *acne*
CC vulgaris, or for stimulating an immune response specific for a *P. acnes*
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against *P. acnes*, or for treating *acne*,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a specifically claimed *P. acnes* polypeptide which is
CC thought to contain an immunogenic region. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 127 AA;
Query Match 93.8%; Score 30; DB 6; Length 127;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
Db 98 IGDIGD 103
RESULT 40
AAU65828
ID AAU65828 standard; protein; 261 AA.
XX AAU65828;
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #26724.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX 02-JUN-2000; 2000US-0208841P.
XX 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59695.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 27023; 1069pp; English.
XX
XX Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC *P. acnes* is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of *P. acnes* in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for *P. acnes* proteins. These antibodies can be used to
CC downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 261 AA;
Query Match 93.8%; Score 30; DB 4; Length 261;
Best Local Similarity 66.7%; Pred. No. 8.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
Db 21 VGDVGD 26
RESULT 41
ABM62347
ID ABM62347 standard; protein; 261 AA.
XX ABM62347;
XX
XX 20-OCT-2003 (first entry)
XX
XX Propionibacterium acnes predicted ORF-encoded polypeptide #27023.
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
XX Propionibacterium acnes.
XX
XX WO2003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skelky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX Barch B, Valliere-Douglass J;
XX WPI; 2003-381789/36.
XX N-PSDB; ACF64624.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a *P. acnes* protein.
XX
XX Example 1; SEQ ID NO 27023; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX encoding a Propionibacterium acnes protein. The invention also relates to
XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX immunogenic fragments of *P. acnes* polypeptides. The invention
XX additionally encompasses expression vectors and host cells comprising a
XX polynucleotide of the invention; antibodies against polypeptides of the
XX invention; fusion proteins comprising a polypeptide of the invention; a

CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

XX Sequence 261 AA;

Query Match 93.8%; Score 30; DB 6; Length 261;
Best Local Similarity 66.7%; Pred. No. 8.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|:|:|
Db 21 VGDVGD 26

RESULT 42
ABG14312

ID ABG14312 standard; protein; 275 AA.

XX AC ABG14312;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #14303.

XX KM Human chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Dermanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS78499.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.

XX XX Claim 20; SEQ ID NO 44671; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptides and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

XX Sequence 275 AA;

Query Match 93.8%; Score 30; DB 4; Length 275;
Best Local Similarity 83.3%; Pred. No. 8.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|:|:|
Db 228 IGDIGD 233

RESULT 43
ABB77414

ID ABB77414 standard; protein; 302 AA.

XX AC ABB77414;

XX DT 29-AUG-2003 (revised)

XX DT 19-JUL-2002 (first entry)

XX DE Fungal membrane transport protein SEQ ID NO 34.

XX KM Fungus; membrane transport; organic acid; amino acid; fatty acid; purine;
XX pyrimidine; nucleotide; nucleoside; lipid; diol; carbohydrate;
XX aromatic compound; vitamin; riboflavin; iodide; food; cosmetic;
XX pharmaceutical; genome mapping; protein structure.

XX OS Eremothecium gossypii.

XX OS Key Location/Qualifiers

XX FT Msc-difference 1..302 /note="Xaa = unknown, encoded by degenerate codons"

XX PN WO200222824-A2.

XX PD 21-MAR-2002.

XX PF 13-SEP-2001; 2001WO-EP010575.

XX PR 15-SEP-2000; 2000DE-01046074.

XX PA (BADT) BASF AG.

XX PI Althoefer H, Revuelta Deval JL, Santos M;

XX DR WPI; 2002-363130/41.

XX DR N-PSDB; ABL58819.

XX PT New nucleic acid from Ashbya gossypii, useful, when modulated, for
XX producing fine chemicals, e.g. amino acids, encodes membrane-construction
XX or -transport proteins.

XX XX Claim 18; Page 161; 193pp; German.

CC The invention relates to an isolated nucleic acid (I) from *Ashbya*
CC *gossypii* comprising: (a) any of the odd-numbered sequences from the fully
CC defined sequences (ABLS8803-ABLS8822) given in the specification; or (b)
CC a fragment of (a). Cells that express (I) show altered production of fine
CC chemicals particularly: (a) organic, amino or fatty acids; (b) purine and
CC pyrimidine bases; (c) nucleotides or nucleosides; (d) diols; (e)
CC carbonyl compounds; (f) aromatic compounds; (g) vitamins (particularly
CC riboflavin); (h) co-factors; and (i) enzymes. These chemicals are useful
CC in the food, fodder, cosmetic and pharmaceutical industries. Also (I) are
CC used for identification and characterisation of *A. gossypii* and related
CC organisms, as primers and probes for detection and amplification of (I).
CC The encoded polypeptides (AB877398-AB877417) are useful as markers for
CC specific gene regions, in genomic mapping and for functional studies on
CC proteins, for evolution or protein structure investigations. Modulating
CC the activity of (I) may improve yields, production and/or efficiency of
CC production of fine chemicals by *A. gossypii*. Particularly the cells show
CC improved transport of waste products from the cell, before these can
CC damage cellular proteins or nucleic acids, so viability is improved.
CC Modulation of (I) may also allow control over the relative amounts of
CC different lipids or fatty acid molecules produced. (Updated on 29-AUG-
CC 2003 to standardise OS field)

XX Sequence 302 AA;

Query Match 93.8%; Score 30; DB 5; Length 302;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 93 LGDIGD 98

RESULT 44

ABP66144 ID ABP66144 standard; protein; 308 AA.

XX AC ABP66144;

XX DT 19-NOV-2002 (first entry)

XX DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:888.

XX Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;

KW antidiarrheic; antibacterial; inhibitor of *Salmonella*; detection;

KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;

XX OS Bifidobacterium longum.

XX PN EP1227152-A1.

XX PD 31-JUL-2002.

XX PF 30-JAN-2001; 2001EP-00102050.

XX PR 30-JAN-2001; 2001EP-00102050.

XX PA (NEST) SOC PROD NESTLE SA.

XX DR WPI; 2002-668397/72.

XX PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as
XX a probe or primer for detecting and/or identifying Bifidobacterium longum
XX in a biological sample.

XX PS Claim 3; SEQ ID NO 888; 80pp; English.

XX The present invention describes a polynucleotide (I) comprising a
XX sequence of a Bifidobacterium genome selected from the nucleotide
XX sequences given in AB081842 and AB081843, or a sequence exhibiting at
XX least 90% identity or which hybridises with the sequences given in
XX AB081842 and AB081843. Also described is a polynucleotide (II) encoding a

CC fusion protein, comprising a sequence selected from 1097 sequences given
CC in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a
CC heterologous polypeptide. (I) has antidiarrheic and antibacterial
CC activities, and can be used as an inhibitor of *Salmonella*. (I) (which is
CC a probe) is useful for the detection and/or identification of
CC Bifidobacterium longum in a biological sample. A carrier containing the
CC lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM I-2618) can be
CC used for preventing and/or treating diarrhoea brought about by pathogenic
CC bacteria and/or rotavirus. The carrier is a food composition selected
CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC products, ice-creams, fermented cereal based products, milk based
CC powders, infant formula, pet food or a pharmaceutical composition
CC selected from tablets, liquid bacterial suspensions, dried oral
CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC (I) is useful in DNA arrays or chips to carry out analysis of the
CC expression of the Bifidobacterium gene. AB081844 to AB081850 represent
CC Bifidobacterium related nucleotide sequences given in the Sequence
CC listing from the present invention but not mentioned further within the
CC specification. N.B. The sequence data for this patent is not represented
CC in the printed specification but is based on sequence information
CC supplied by the European Patent Office

XX Sequence 308 AA;

Query Match 93.8%; Score 30; DB 5; Length 308;
Best Local Similarity 66.7%; Pred. No. 9.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 171 VGDVGD 176

RESULT 45

ABM65924 ID ABM65924 standard; protein; 353 AA.

XX AC ABM65924;

XX DT 20-OCT-2003 (first entry)

XX DE Propionibacterium acnes immunogenic polypeptide #30600.

KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine; immunogenic.

XX OS Propionibacterium acnes.

XX PN WO2003033515-A1.

XX PD 24-APR-2003.

XX PF 11-OCT-2002; 2002WO-US032727.

XX PR 15-OCT-2001; 2001US-00978825.

XX PA (CORI-) CORIXA CORP.

XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JT,
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX Barch B, Valliere-Douglass J;

XX DR WPI; 2003-381789/36.

XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a P. acnes protein.

XX PS Claim 7; SEQ ID NO 30600; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX encoding a Propionibacterium acnes protein. The invention also relates to
XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of *P. acnes* polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a *P. acnes*
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising *P. acnes* polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of *P. acnes* in a
CC patient; and a method for inhibiting the development of *P. acnes* in a
CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating *acne*
CC vulgaris, or for stimulating an immune response specific for a *P. acnes*
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against *P. acnes*, or for treating *acne*,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a specifically claimed *P. acnes* polypeptide which is
CC thought to contain an immunogenic region. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 353 AA;

Query Match 93.8%; Score 30; DB 6; Length 353;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|:|:
Db 117 VGDVGD 122

RESULT 46

ABO71148 standard; protein; 456 AA.

AC ABO71148;

DT 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #3323.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD04719.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 19894; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the

CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences ABO67826-
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

SQ Sequence 456 AA;

Query Match 93.8%; Score 30; DB 7; Length 456;
Best Local Similarity 83.3%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|:|:
Db 410 IGDIGD 415

RESULT 47

ADQ36971 standard; protein; 500 AA.

AC ADQ36971;

DT 07-OCT-2004 (first entry)

XX Cell proliferation-related polypeptide #63.

XX cell proliferation related polypeptide; cell proliferation; senescence;

KM differentiation; stress response.

OS Oryza sativa.

PN WO2004061122-A2.

PD 22-JUL-2004.

PF 23-DEC-2003; 2003WO-US041200.

PR 26-DEC-2002; 2002US-0436565P.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Cooper B;

DR WPI; 2004-534388/51.

XX New nucleic acid molecule encoding a cell proliferation-related
XX polypeptide, useful for modulating cell proliferation, senescence,
XX differentiation, development, and stress response in plants, and for
XX producing enhanced food crops.

PS Claim 28; SEQ ID NO 130; 408pp; English.

XX The present invention relates to an isolated nucleic acid molecule
XX encoding a cell proliferation-related polypeptide. The nucleic acid
XX molecule and the encoded polypeptide, and methods are useful for
XX modulating cell proliferation, senescence, differentiation, development,
XX and stress response in plants, and for producing enhanced food crops. The
XX present sequence represents a cell proliferation-related polypeptide. The
XX specification as EPO data.

SQ Sequence 500 AA;

Query Match 93.8%; Score 30; DB 8; Length 500;
 Best Local Similarity 66.7%; Pred. No. 1.6e+03;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
 :||:|
 Db 115 VGDVGD 120

RESULT 48
 ADQ15641
 ID ADQ15641 standard; protein; 500 AA.

XX ADQ15641;
 XX 07-OCT-2004 (first entry)
 XX
 XX Rice stress-related protein #25.

XX rice; stress-related protein; plant maturation; plant development;
 XX plant proliferation; plant senescence; plant disease-resistance;
 XX plant stress response; transgenic plant; pest tolerance;
 XX herbicide tolerance; biotic stress tolerance; abiotic stress tolerance;
 XX improved nutritional value; increased yield; increased proliferation.

XX Oryza sativa.
 XX WO2004061080-A2.

XX 22-JUL-2004.

XX 23-DEC-2003; 2003WO-US041098.

XX 26-DEC-2002; 2002US-0436564P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Cooper B;

XX WPI; 2004-534374/51.

XX N-PSDB; ADQ15640.

XX New isolated nucleic acids and proteins, useful for producing transgenic
 PT plants having improved properties, e.g. tolerance to pests, herbicides,
 PT or biotic or abiotic stresses, improved nutritional value, or increased
 PT yield or proliferation.
 XX

PS Claim 28; SEQ ID NO 50; 551bp; English.

XX The invention comprises the amino acid and coding sequences of rice
 CC stress-related proteins. The DNA and protein sequences of the invention
 CC are useful for regulating and controlling plant maturation and
 CC development, including proliferation, senescence, disease-resistance, or
 CC stress response. They are also useful for producing transgenic plants
 CC having improved properties, e.g. tolerance to pests, herbicides, or
 CC biotic or abiotic stresses, improved nutritional value, increased yield
 CC or proliferation, or improved structure causing less loss from lodging or
 CC shattering. The present amino acid sequence represents a rice stress-
 CC related protein of the invention.

XX Sequence 500 AA;

Qy Query Match 93.8%; Score 30; DB 8; Length 500;
 Best Local Similarity 66.7%; Pred. No. 1.6e+03;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
 :||:|
 Db 115 VGDVGD 120

RESULT 49

ADQ15026
 ID ADQ15026 standard; protein; 505 AA.
 XX
 AC ADQ15026;

XX 23-SEP-2004 (first entry)

XX Synchocystis sp. PCC 6803 amylo maltase polypeptide.

XX Amylo maltase; transgenic; foodstuff; food additive; modifier;
 XX saccharide; (alpha)-1,4-glucan; cooked rice; Japanese confectionery;
 XX snack; noodle; Chinese dumpling; shao-mai; fishery kneading;
 XX refrigerated; baby food; pet; animal feed; drink; food supplement;
 XX cyclic glucan; Synchocystis sp. PCC 6803.

XX Synchocystis sp. PCC 6803.

XX JP2004187674-A.

XX 08-JUL-2004.

XX 12-SEP-2003; 2003JP-00322319.

XX 26-NOV-2002; 2002JP-00342966.

XX (EZAKI) EZAKI GLICO CO LTD.

XX (NIDE) NIDE CORP.

XX WPI; 2004-528719/51.

XX N-PSDB; ADQ15025.

XX New mutated amylo maltase polypeptide having increased enzyme activity
 PT and reduced hydrolyzing activity, useful for manufacturing foodstuffs
 PT such as Japanese confectionery, noodles, baby foods, and food additives.
 XX

PS Example 9; SEQ ID NO 17; 51bp; Japanese.

XX The invention relates to a novel amylo maltase polypeptide comprising the
 CC amino acid sequence of a wild-type amylo maltase with a substitution,
 CC addition or deletion at a position, where the amino acid residue
 CC interacts with acarbose of amino acids other than the wild-type amylo
 CC maltase amino acid sequence. The invention further comprises: a nucleic
 CC acid molecule containing a nucleic acid sequence encoding the novel amylo
 CC maltase polypeptide; a vector containing the nucleic acid sequence; a
 CC cell containing the nucleic acid sequence; biological tissue containing
 CC the nucleic acid sequence; a transgenic organism containing the nucleic
 CC acid sequence; a foodstuff, a food additive or a modifier of foodstuff
 CC containing the nucleic acid sequence; a computer readable recording
 CC medium which contains the information of the nucleic acid sequence
 CC encoding the amino acid sequence of the novel amylo maltase polypeptide;
 CC and a cyclic glucan obtained by reacting the novel amylo maltase
 CC polypeptide on saccharides which have a linear structure of (alpha)-1,4-
 CC glucan. The novel amylo maltase polypeptide is useful for manufacturing
 CC foodstuffs such as cooked rice, Japanese confectionery, snacks, bakeries,
 CC noodles, Chinese dumpling, shao-mai, fishery kneading goods,
 CC refrigerating foods, baby foods, pet foods, feed for animals, drinks, and
 CC food supplements, where the method involves adding the novel amylo
 CC maltase polypeptide to the foodstuff before or immediately after heat
 CC processing of the raw material, where the novel amylo maltase polypeptide
 CC generates a cyclic glucan from the starch of the foodstuff. The novel
 CC amylo maltase polypeptide is useful for manufacturing glucan which has a
 CC cyclic structure of (alpha)-1,4-glycoside linkage, foodstuff materials
 CC and food additives, modifier of foodstuffs, a drink or eatable
 CC composition, infusion solution or composition for adhesion, where the
 CC method involves reacting the saccharide containing the linear structure
 CC of (alpha)-1,4-glycoside linkage or its derivatives with the novel amylo
 CC maltase polypeptide. This sequence represents a Synchocystis sp. PCC
 CC 6803 amylo maltase polypeptide of the invention.

XX Sequence 505 AA;

Qy Query Match 93.8%; Score 30; DB 8; Length 505;
 Best Local Similarity 83.3%; Pred. No. 1.6e+03;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IGDIGD 6
 |||:
 22 IGDIGD 27

RESULT 50

ABO64512

ID ABO64512 standard; protein; 544 AA.

AC ABO64512;

DT 29-JUL-2004 (first entry)

DE Klebsiella pneumoniae polypeptide seqid 11029.

KW Recombinant expression vector; transcription regulatory element;

XX Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX Klebsiella pneumoniae.

OS US6610836-B1.

PD 26-AUG-2003.

PF 27-JAN-2000; 2000US-00489039.

PR 29-JAN-1999; 99US-0117747P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Breton GL, Osborne M;

DR WPI: 2003-895346/82.

DR N-PSDB; ACH98063.

PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 preparing a vaccine composition against Klebsiella pneumoniae.

PS Disclosure; SEQ ID NO 11029; 932pp; English.

CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention

SQ Sequence 544 AA;

Query Match 93.8%; Score 30; DB 7; Length 544;
 Best Local Similarity 66.7%; Pred. NO. 1.7e+03;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :||:
 Db 121 VGDVGD 126

Search completed: July 27, 2005, 01:17:29
 Job time : 167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 01:13:14 ; Search time 22 Seconds

(without alignments)
20.359 Million cell updates/sec

Title: US-10-027-015A-4
Perfect score: 32
Sequence: 1 IGDIGD 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	325	4	US-09-248-796A-20030
2	31	96.9	59	1	US-08-470-179-7
3	31	96.9	69	4	US-09-513-999C-5488
4	31	96.9	86	4	US-09-513-999C-5489
5	31	96.9	185	4	US-09-902-540-14654
6	31	96.9	209	4	US-09-248-796A-21116
7	31	96.9	211	1	US-08-631-607-4
8	31	96.9	211	4	US-09-098-358B-4
9	31	96.9	1971	4	US-09-914-272A-1
10	31	96.9	1971	4	US-10-638-333-1
11	30	93.8	365	4	US-09-902-540-16239
12	30	93.8	423	4	US-09-902-540-10934
13	30	93.8	456	4	US-09-252-991A-19894
14	30	93.8	544	4	US-09-489-039A-11029
15	30	93.8	831	4	US-09-252-991A-30097
16	30	93.8	1257	4	US-09-252-991A-17290
17	29	90.6	443	4	US-09-552-991A-20035
18	29	90.6	447	4	US-09-252-991A-26752
19	29	90.6	514	4	US-09-252-991A-30600
20	29	90.6	674	4	US-09-252-991A-26864
21	29	90.6	737	4	US-09-902-540-15691
22	29	90.6	779	4	US-09-252-991A-27761
23	28	87.5	72	4	US-09-489-039A-7652
24	28	87.5	96	4	US-09-270-767-41400
25	28	87.5	96	4	US-09-270-767-56616
26	28	87.5	137	4	US-09-583-110-4901
27	28	87.5	142	4	US-09-107-433-5189

28	28	87.5	146	4	US-09-328-352-7997	Sequence 7997, Ap
29	28	87.5	151	4	US-09-489-039A-7930	Sequence 7930, Ap
30	28	87.5	198	4	US-09-710-279-2188	Sequence 2188, Ap
31	28	87.5	223	4	US-09-710-279-1234	Sequence 1234, Ap
32	28	87.5	226	4	US-09-902-540-10537	Sequence 10537, A
33	28	87.5	250	4	US-09-614-912-62	Sequence 62, Appl
34	28	87.5	264	4	US-09-614-912-74	Sequence 74, Appl
35	28	87.5	337	4	US-09-710-279-148	Sequence 148, App
36	28	87.5	337	4	US-09-710-279-156	Sequence 156, App
37	28	87.5	337	4	US-09-710-279-458	Sequence 458, App
38	28	87.5	337	4	US-09-710-279-1152	Sequence 1152, Ap
39	28	87.5	337	4	US-09-710-279-1198	Sequence 1198, Ap
40	28	87.5	337	4	US-09-710-279-1220	Sequence 1220, Ap
41	28	87.5	337	4	US-09-710-279-1384	Sequence 1384, Ap
42	28	87.5	337	4	US-09-710-279-2044	Sequence 2044, Ap
43	28	87.5	337	4	US-09-710-279-2196	Sequence 2196, Ap
44	28	87.5	337	4	US-09-710-279-2785	Sequence 2786, Ap
45	28	87.5	353	4	US-09-252-991A-23322	Sequence 23322, A
46	28	87.5	361	4	US-09-198-452A-812	Sequence 812, App
47	28	87.5	372	4	US-09-614-912-82	Sequence 82, Appl
48	28	87.5	375	3	US-09-134-001C-3918	Sequence 3918, Ap
49	28	87.5	380	4	US-09-438-185A-764	Sequence 764, App
50	28	87.5	388	4	US-09-583-110-3650	Sequence 3650, Ap

ALIGNMENTS

RESULT 1
US-09-248-796A-20030
; Patent No. 6747137
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.1132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20030
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20030
Query Match 100.0%; Score 32; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
DB 299 IGDIGD 304
RESULT 2
US-08-470-179-7
; Sequence 7, Application US/08470179
; Patent No. 5645994
; GENERAL INFORMATION:
; APPLICANT: Huang Ph.D, Wai Mun
; TITLE OF INVENTION: Method and Compositions for
; NUMBER OF SEQUENCES: Identification of Species in a Sample
; CORRESPONDENCE ADDRESS: 207
; ADDRESSEE: Traak, Bitit and Roosa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah

```

; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,179
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Ph.D, Susan E.
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 801-532-1922
; TELEFAX: 801-531-9168
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 59 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Rickettsiae rickettsii
US-08-470-179-7
```

```

Query Match          96.9%; Score 31; DB 1; Length 59;
Best Local Similarity 83.3%; Pred. No. 22;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 IGDIGD 6
       :|||||
Db      7 IGDVGD 12
```

```

RESULT 3
US-09-513-999C-5488
; Sequence 5488, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5488
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5488
```

```

Query Match          96.9%; Score 31; DB 4; Length 69;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 IGDIGD 6
       :|||||
Db      40 VGDIGD 45
```

```

RESULT 4
US-09-513-999C-5489
; Sequence 5489, Application US/09513999C
```

```

; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5489
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5489
```

```

Query Match          96.9%; Score 31; DB 4; Length 86;
Best Local Similarity 83.3%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 IGDIGD 6
       :|||||
Db      57 VGDIGD 62
```

```

RESULT 5
US-09-902-540-14654
; Sequence 14654, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14654
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(185)
; OTHER INFORMATION: unsure at all Xaa locations
US-09-902-540-14654
```

```

Query Match          96.9%; Score 31; DB 4; Length 185;
Best Local Similarity 83.3%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 IGDIGD 6
       :|||||
Db      62 VGDIGD 67
```

```

RESULT 6
US-09-248-796A-21116
; Sequence 21116, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
```

CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 21116
LENGTH: 209
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-21116

Query Match 96.9%; Score 31; DB 4; Length 209;
Best Local Similarity 83.3%; Pred. No. 84;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 1 IGDIGD 6

RESULT 7
US-08-631-607-4
Sequence 4, Application US/08631607
Patent No. 5767252

GENERAL INFORMATION:
APPLICANT: Worley et al, Paul
TITLE OF INVENTION: NOVEL NEURONAL CELL GROWTH FACTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/631,607
FILING DATE: 08-APR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/086001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-631-607-4

Query Match 96.9%; Score 31; DB 1; Length 211;
Best Local Similarity 83.3%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 151 VGDIGD 156

RESULT 8
US-09-098-358B-4

Sequence 4, Application US/09098358B
Patent No. 6436673
GENERAL INFORMATION:
APPLICANT: THE JOHN HOPKINS SCHOOL OF MEDICINE
APPLICANT: WORLEY, PAUL
TITLE OF INVENTION: NOVEL NEURONAL CELL GROWTH FACTOR
FILE REFERENCE: JHU 1340-1 (07265/086002)
CURRENT APPLICATION NUMBER: US/09/098,358B
CURRENT FILING DATE: 1998-06-16
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent version 3.0
SEQ ID NO 4
LENGTH: 211
TYPE: PRT
ORGANISM: ARTIFICIAL
FEATURE:
OTHER INFORMATION: RAT CRP
US-09-098-358B-4

Query Match 96.9%; Score 31; DB 4; Length 211;
Best Local Similarity 83.3%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 151 VGDIGD 156

RESULT 9
US-09-914-272A-1
Sequence 1, Application US/09914272A
Patent No. 6673913
GENERAL INFORMATION:
APPLICANT: Sakaguchi, No. 6673913uo
APPLICANT: Kuwahara, Kazuhiko
TITLE OF INVENTION: GANP Protein
FILE REFERENCE: 050208-0014
CURRENT APPLICATION NUMBER: US/09/914,272A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: PCT/JP99/04634
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: 47035/1999
PRIOR FILING DATE: 1999-02-24
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent version 3.1
SEQ ID NO 1
LENGTH: 1971
TYPE: PRT
ORGANISM: Mouse
US-09-914-272A-1

Query Match 96.9%; Score 31; DB 4; Length 1971;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 1388 VGDIGD 1393

RESULT 10
US-10-638-333-1
Sequence 1, Application US/10638333
Patent No. 6825020
GENERAL INFORMATION:
APPLICANT: Sakaguchi, No. 6825020uo
APPLICANT: Kuwahara, Kazuhiko
TITLE OF INVENTION: GANP Protein
FILE REFERENCE: 050208-0014
CURRENT APPLICATION NUMBER: US/10/638,333
CURRENT FILING DATE: 2003-08-12
PRIOR APPLICATION NUMBER: PCT/JP99/04634

;; PRIOR FILING DATE: 1999-08-27
;; PRIOR APPLICATION NUMBER: 47035/1999
;; PRIOR FILING DATE: 1999-02-24
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 1971
;; TYPE: PRT
;; ORGANISM: Mouse
US-10-638-333-1

Query Match 96.9%; Score 31; DB 4; Length 1971;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 1388 VGDVGD 1393

RESULT 11
US-09-902-540-16239
; Sequence 16239, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(11849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16239
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16239

Query Match 93.8%; Score 30; DB 4; Length 365;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 150 VGDVGD 155

RESULT 12
US-09-902-540-10934
; Sequence 10934, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(11849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10934
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10934

Query Match 93.8%; Score 30; DB 4; Length 423;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 128 LGDIGD 133

RESULT 13
US-09-252-991A-19894
; Sequence 19894, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19894
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19894

Query Match 93.8%; Score 30; DB 4; Length 456;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 410 IGDIGD 415

RESULT 14
US-09-489-039A-11029
; Sequence 11029, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11029
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11029

Query Match 93.8%; Score 30; DB 4; Length 544;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 121 VGDVGD 126

RESULT 15
US-09-252-991A-30097
; Sequence 30097, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30097
LENGTH: 831
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30097

Query Match 93.8%; Score 30; DB 4; Length 831;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
: ||: ||
Db 238 VGDVGD 243

RESULT 16
US-09-252-991A-17290
Sequence 17290, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17290
LENGTH: 1257
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17290

Query Match 93.8%; Score 30; DB 4; Length 1257;
Best Local Similarity 83.3%; Pred. No. 9.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
: ||: ||
Db 423 LGDIGD 428

RESULT 17
US-09-252-991A-20035
Sequence 20035, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20035
LENGTH: 443
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20035

Query Match 90.6%; Score 29; DB 4; Length 443;
Best Local Similarity 66.7%; Pred. No. 4.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
: ||: ||
Db 303 LGDVGD 308

RESULT 18
US-09-252-991A-26752
Sequence 26752, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26752
LENGTH: 447
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26752

Query Match 90.6%; Score 29; DB 4; Length 447;
Best Local Similarity 66.7%; Pred. No. 4.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
: ||: ||
Db 158 LGDVGD 163

RESULT 19
US-09-252-991A-30600
Sequence 30600, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30600
LENGTH: 514
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30600

Query Match 90.6%; Score 29; DB 4; Length 514;
Best Local Similarity 66.7%; Pred. No. 5.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|:|:
Db 216 IGDVGD 221

RESULT 20

US-09-252-991A-26864
; Sequence 26864, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26864
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (18),(19)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-26864

Query Match 90.6%; Score 29; DB 4; Length 674;
Best Local Similarity 66.7%; Pred. No. 7.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|:|:
Db 312 VGDLGD 317

RESULT 21

US-09-902-540-15691
; Sequence 15691, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15691
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15691

Query Match 90.6%; Score 29; DB 4; Length 677;
Best Local Similarity 66.7%; Pred. No. 7.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|:|:
Db 479 IGDVGD 484

RESULT 22

US-09-252-991A-27761

; Sequence 27761, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27761
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (739)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-27761

Query Match 90.6%; Score 29; DB 4; Length 739;
Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|:|:
Db 667 VGDLGD 672

RESULT 23
US-09-489-039A-7652
; Sequence 7652, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7652
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7652

Query Match 87.5%; Score 28; DB 4; Length 72;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
|||:
Db 49 GDIGD 53

RESULT 24
US-09-270-767-41400
; Sequence 41400, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 41400
LENGTH: 96
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41400

Query Match 87.5%; Score 28; DB 4; Length 96;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
DB 33 IGDIGD 38

RESULT 25
US-09-270-767-56616
Sequence 56616, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 56616
LENGTH: 96
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-56616

Query Match 87.5%; Score 28; DB 4; Length 96;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
DB 33 IGDIGD 38

RESULT 26
US-09-583-110-4901
Sequence 4901, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 4901
LENGTH: 137
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-4901

Query Match 87.5%; Score 28; DB 4; Length 137;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
DB 52 IGDIGD 57

RESULT 27
US-09-107-433-5189
Sequence 5189, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Maltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5189:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...142
SEQUENCE DESCRIPTION: SEQ ID NO: 5189:
US-09-107-433-5189

Query Match 87.5%; Score 28; DB 4; Length 142;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
DB 57 IGDIGD 62

RESULT 28
US-09-328-352-7997
Sequence 7997, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAWANNIT FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7997
LENGTH: 146
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7997

Query Match 87.5%; Score 28; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
Db 33 GDIGD 37

RESULT 29
US-09-489-039A-7930
Sequence 7930, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7930
LENGTH: 151
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7930

Query Match 87.5%; Score 28; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
Db 77 GDIGD 81

RESULT 30
US-09-710-279-2188
Sequence 2188, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P13480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 2188
LENGTH: 198
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-2188

Query Match 87.5%; Score 28; DB 4; Length 198;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 83 IGDIGD 88

RESULT 31
US-09-710-279-1234
Sequence 1234, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P13480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 1234
LENGTH: 223
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1234

Query Match 87.5%; Score 28; DB 4; Length 223;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 108 IGDIGD 113

RESULT 32
US-09-902-540-10537
Sequence 10537, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 10537
LENGTH: 226
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-10537

Query Match 87.5%; Score 28; DB 4; Length 226;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 40 IGDIGD 45

```
RESULT 33
US-09-614-912-62
; Sequence 62, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Cai, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/172,959
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/172,946
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 62
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (123)
US-09-614-912-62

Query Match          87.5%; Score 28; DB 4; Length 250;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 34
US-09-614-912-74
; Sequence 74, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Cai, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412

RESULT 35
US-09-710-279-148
; Sequence 148, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 148
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-148

Query Match          87.5%; Score 28; DB 4; Length 337;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 36
US-09-710-279-156
; Sequence 156, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 156
```

LENGTH: 337
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-156

Query Match 87.5%; Score 28; DB 4; Length 337;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 222 IGEIGD 227

RESULT 37
US-09-710-279-458
Sequence 458, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 458
LENGTH: 337
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-458

Query Match 87.5%; Score 28; DB 4; Length 337;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 222 IGEIGD 227

RESULT 38
US-09-710-279-1152
Sequence 1152, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1152
LENGTH: 337
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1152

Query Match 87.5%; Score 28; DB 4; Length 337;

Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 222 IGEIGD 227

RESULT 39
US-09-710-279-1198
Sequence 1198, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1198
LENGTH: 337
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1198

Query Match 87.5%; Score 28; DB 4; Length 337;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 222 IGEIGD 227

RESULT 40
US-09-710-279-1220
Sequence 1220, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1220
LENGTH: 337
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1220

Query Match 87.5%; Score 28; DB 4; Length 337;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 222 IGEIGD 227

RESULT 41

US-09-710-279-1384
; Sequence 1384, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1384
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1384

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 337;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
||:||||
DB 222 IGEIGD 227

RESULT 42
US-09-710-279-2044
; Sequence 2044, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2044
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2044

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 337;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
||:||||
DB 222 IGEIGD 227

RESULT 43
US-09-710-279-2196
; Sequence 2196, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09

; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2196
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2196

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 337;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
||:||||
DB 222 IGEIGD 227

RESULT 44
US-09-710-279-2786
; Sequence 2786, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2786
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2786

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 337;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
||:||||
DB 222 IGEIGD 227

RESULT 45
US-09-252-991A-23322
; Sequence 23322, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23322
; LENGTH: 353
; TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23322

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 353;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
Db 58 GDIGD 62

RESULT 46
US-09-198-452A-812
Sequence 812, Application US/09198452A

PATENT No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention
TITLE OF INVENTION: and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT FILING DATE: US/09/198,452A
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 812
LENGTH: 361
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-812

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 361;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
Db 166 GDIGD 170

RESULT 47
US-09-614-912-82
Sequence 82, Application US/09614912
PATENT No. 6677502
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalski, Antoni
APPLICANT: Orozco, Buddy
APPLICANT: Miao, Gou-Hau
APPLICANT: Famodu, Omolayo O.
APPLICANT: Lee, Jian Ming
APPLICANT: Sakai, Hajime
APPLICANT: Weng, Zude
APPLICANT: Cai, Perry G
APPLICANT: Anderson, Shawn
TITLE OF INVENTION: Plant Metabolism Genes
FILE REFERENCE: B1378 US NA
CURRENT APPLICATION NUMBER: US/09/614,912
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 60/143,401
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/143,412
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/146,650
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/170,906
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: 60/172,959
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/172,946
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 204
SOFTWARE: Microsoft Office 97

SEQ ID NO 82
LENGTH: 372
TYPE: PRT
ORGANISM: Oryza sativa
US-09-614-912-82

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 372;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
Db 186 GDIGD 190

RESULT 48
US-09-134-001C-3918
Sequence 3918, Application US/09134001C
PATENT No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3918
LENGTH: 375
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3918

Query Match
Best Local Similarity 87.5%; Score 28; DB 3; Length 375;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 260 IGDIGD 265

RESULT 49
US-09-438-185A-764
Sequence 764, Application US/09438185A
PATENT No. 6822071
GENERAL INFORMATION:
APPLICANT: Stephens, Richard
APPLICANT: Mitchell, Wayne
APPLICANT: Kalman, Sue
APPLICANT: Davis, Ronald
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
FILE REFERENCE: 018941-000411US
CURRENT APPLICATION NUMBER: US/09/438,185A
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/128,606
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 1074
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 764
LENGTH: 380
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
FEATURE:
OTHER INFORMATION: CPN0762
US-09-438-185A-764

Query Match 87.5%; Score 28; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 6,2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
|||
Db 185 GDIGD 189

RESULT 50
US-09-583-110-3650
; Sequence 3650, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3650
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3650

Query Match 87.5%; Score 28; DB 4; Length 388;
Best Local Similarity 100.0%; Pred. No. 6,4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
|||
Db 170 GDIGD 174

Search completed: July 27, 2005, 01:21:40
Job time : 24 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 01:15:14 ; Search time 153 Seconds

(without alignments)
15.255 Million cell updates/sec

Title: US-10-027-015A-4

Perfect score: 32

Sequence: 1 IGDIGD 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1741741 seqs, 388992284 residues

Total number of hits satisfying chosen parameters: 1741741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	32	100.0	6	18 US-10-027-015A-4	Sequence 3, Appl1
2	32	100.0	6	18 US-10-027-015A-4	Sequence 4, Appl1
3	32	100.0	153	15 US-10-767-701-53899	Sequence 53899, A
4	32	100.0	515	15 US-10-369-493-18184	Sequence 18184, A
5	32	100.0	689	16 US-10-425-115-12813	Sequence 192813, A
6	31	96.9	80	16 US-10-767-701-36746	Sequence 36746, A
7	31	96.9	81	14 US-10-029-386-29095	Sequence 29095, A
8	31	96.9	126	16 US-10-425-115-238465	Sequence 238465, A
9	31	96.9	129	15 US-10-158-034-94	Sequence 94, Appl1
10	31	96.9	156	15 US-10-158-034-68	Sequence 68, Appl1
11	31	96.9	166	14 US-10-205-342-21	Sequence 21, Appl1

12	31	96.9	166	14	US-10-205-823-72	Sequence 72, Appl1
13	31	96.9	166	16	US-10-408-765A-1910	Sequence 1910, Ap
14	31	96.9	176	16	US-10-425-115-256293	Sequence 256293, A
15	31	96.9	205	15	US-10-158-034-93	Sequence 93, Appl1
16	31	96.9	211	14	US-10-324-951-4	Sequence 4, Appl1
17	31	96.9	216	16	US-10-739-930-7500	Sequence 7500, Ap
18	31	96.9	229	16	US-10-739-930-7499	Sequence 7499, Ap
19	31	96.9	328	14	US-10-156-761-11418	Sequence 11418, A
20	31	96.9	342	15	US-10-389-566-367	Sequence 367, App
21	31	96.9	364	15	US-10-369-493-23444	Sequence 23444, A
22	31	96.9	464	15	US-10-424-599-218955	Sequence 218355, A
23	31	96.9	537	16	US-10-425-115-254563	Sequence 254563, A
24	31	96.9	693	15	US-10-369-493-11876	Sequence 11876, A
25	31	96.9	1971	16	US-10-747-133-1	Sequence 1, Appl1
26	30	93.8	115	15	US-10-424-599-235305	Sequence 235305, A
27	30	93.8	115	17	US-10-732-923-4243	Sequence 4243, Ap
28	30	93.8	204	16	US-10-425-115-185817	Sequence 185817, A
29	30	93.8	247	16	US-10-425-115-335683	Sequence 335683, A
30	30	93.8	328	15	US-10-425-114-52030	Sequence 52030, A
31	30	93.8	477	15	US-10-425-114-68337	Sequence 68337, A
32	30	93.8	486	15	US-10-425-114-57203	Sequence 57203, A
33	30	93.8	486	16	US-10-425-115-325328	Sequence 325328, A
34	30	93.8	540	16	US-10-425-115-239897	Sequence 239897, A
35	30	93.8	567	15	US-10-424-599-170978	Sequence 170978, A
36	30	93.8	568	16	US-10-437-963-184415	Sequence 184405, A
37	30	93.8	598	16	US-10-437-963-139894	Sequence 139894, A
38	30	93.8	664	16	US-10-437-963-181573	Sequence 181573, A
39	30	93.8	3415	17	US-10-937-730A-6	Sequence 6, Appl1
40	29	90.6	73	15	US-10-424-599-227900	Sequence 227900, A
41	29	90.6	79	16	US-10-425-115-253371	Sequence 253371, A
42	29	90.6	90	15	US-10-424-599-263936	Sequence 263936, A
43	29	90.6	152	16	US-10-425-115-355214	Sequence 355214, A
44	29	90.6	174	16	US-10-767-701-60220	Sequence 60220, A
45	29	90.6	187	16	US-10-437-963-146654	Sequence 146654, A
46	29	90.6	189	16	US-10-425-115-201329	Sequence 201329, A
47	29	90.6	321	16	US-10-437-963-156021	Sequence 156021, A
48	29	90.6	355	16	US-10-437-963-102521	Sequence 102521, A
49	29	90.6	393	14	US-10-156-761-146659	Sequence 146659, A
50	29	90.6	422	15	US-10-425-114-45513	Sequence 45513, A

ALIGNMENTS

RESULT 1
US-10-027-015A-3
Sequence 3, Application US/10027015A
Publication No. US20050143312A1
GENERAL INFORMATION:
APPLICANT: AKELLA, RAMA
APPLICANT: RANIERI, JOHN P.
TITLE OF INVENTION: COMBINATIONS AND METHODS FOR PROMOTING MYOCARDIAL AND PERIPHERAL
FILE REFERENCE: 2103.000500
CURRENT PILING DATE: 2001-12-21
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: SYNTHETIC PEPTIDE
US-10-027-015A-3
Query Match
Best Local Similarity 100.0%; Score 32; DB 18; Length 6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IGDIGD 6
Db 1 IGDIGD 6

RESULT 2
 US-10-027-015A-4
 ; Sequence 4, Application US/10027015A
 ; Publication No. US20050143312A1
 ; GENERAL INFORMATION:
 ; APPLICANT: AKELLA, RAMA
 ; APPLICANT: RANIERI, JOHN P.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROMOTING MYOCARDIAL AND PERIPHERAL
 ; TITLE OF INVENTION: ANGIOGENESIS
 ; FILE REFERENCE: 2103.000500
 ; CURRENT APPLICATION NUMBER: US/10/027,015A
 ; CURRENT FILING DATE: 2001-12-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4
 ; LENGTH: 6
 ; TYPE: PRT
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: SYNTHETIC PEPTIDE
 US-10-027-015A-4

Query Match 100.0%; Score 32; DB 16; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 Db 1 IGDIGD 6

RESULT 3
 US-10-767-701-53899
 ; Sequence 53899, Application US/10767701
 ; Publication No. US20040172684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-29
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 53899
 ; LENGTH: 153
 ; TYPE: PRT
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 13784568.pep
 US-10-767-701-53899

Query Match 100.0%; Score 32; DB 16; Length 153;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 Db 33 IGDIGD 38

RESULT 4
 US-10-369-493-18184
 ; Sequence 18184, Application US/10369493
 ; Publication No. US2003023675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 18184
 ; LENGTH: 515
 ; TYPE: PRT
 ; ORGANISM: Thermoplasma acidophilum
 US-10-369-493-18184

Query Match 100.0%; Score 32; DB 15; Length 515;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 Db 423 IGDIGD 428

RESULT 5
 US-10-425-115-192813
 ; Sequence 192813, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 192813
 ; LENGTH: 689
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1) .. (689)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: M714577_107432C.1.pep
 US-10-425-115-192813

Query Match 100.0%; Score 32; DB 16; Length 689;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 Db 423 IGDIGD 428

RESULT 6
 US-10-767-701-36746
 ; Sequence 36746, Application US/10767701
 ; Publication No. US20040172684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-29

```

; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 36746
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1053_1.pep
US-10-767-701-36746

Query Match          96.9%; Score 31; DB 16; Length 80;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :|||||
DB      37 VGDIGD 42

RESULT 7
US-10-029-386-29095
; Sequence 29095, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29095
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
US-10-029-386-29095

Query Match          96.9%; Score 31; DB 14; Length 81;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :|||||
DB      40 VGDIGD 45

RESULT 8
US-10-425-115-238465
; Sequence 238465, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 238465
```

```

; LENGTH: 126
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_149069C.1.pep
US-10-425-115-238465

Query Match          96.9%; Score 31; DB 16; Length 126;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :|||||
DB      65 IGDVGD 70

RESULT 9
US-10-158-034-94
; Sequence 94, Application US/10158034
; Publication No. US20030219758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT210C1
; CURRENT APPLICATION NUMBER: US/10/158,034
; CURRENT FILING DATE: 2002-05-31
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 94
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (123)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (128)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-158-034-94

Query Match          96.9%; Score 31; DB 15; Length 129;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :|||||
DB      69 VGDIGD 74

RESULT 10
US-10-158-034-68
; Sequence 68, Application US/10158034
; Publication No. US20030219758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT210C1
; CURRENT APPLICATION NUMBER: US/10/158,034
```

```

; CURRENT FILING DATE: 2002-05-31
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-034-68

Query Match
Best Local Similarity 96.9%; Score 31; DB 15; Length 156;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 101 VGDIGD 106

RESULT 11
US-10-205-342-21
; Sequence 21, Application US/10205342
; Publication No. US20030108906A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Plincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WI-A-018198
; CURRENT APPLICATION NUMBER: US/10/205,342
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Version 3.0
; SEQ ID NO 21
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Mus musculus
; OTHER INFORMATION: Protein: Cofilin
US-10-205-342-21

Query Match
Best Local Similarity 96.9%; Score 31; DB 14; Length 166;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 57 VGDIGD 62

RESULT 12
US-10-205-823-72
; Sequence 72, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF PROSTATE CANCER
```

```

; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-72

Query Match
Best Local Similarity 96.9%; Score 31; DB 14; Length 166;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 57 VGDIGD 62

RESULT 13
US-10-408-765A-1910
; Sequence 1910, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1910
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1910

Query Match
Best Local Similarity 96.9%; Score 31; DB 16; Length 166;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 57 VGDIGD 62

RESULT 14
US-10-425-115-256293
; Sequence 256293, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
```

;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;; TITLE OF INVENTION: Plants
;; FILE REFERENCE: 38-21(53222)B
;; CURRENT APPLICATION NUMBER: US/10/425,115
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 369326
;; SEQ ID NO 256293
;; LENGTH: 176
;; TYPE: PRT
;; ORGANISM: Zea mays
;; FEATURE:
;; OTHER INFORMATION: Clone ID: MRT4577_165330C.1.pdp
US-10-425-115-256293

Query Match 96.9%; Score 31; DB 16; Length 176;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
Db 57 VGDIGD 62

RESULT 15
US-10-158-034-93
; Sequence 93, Application US/10158034
; Publication No. US20030219758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT210C1
; CURRENT APPLICATION NUMBER: US/10/158,034
; CURRENT FILING DATE: 2002-05-31
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 93
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-034-93

Query Match 96.9%; Score 31; DB 15; Length 205;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
Db 96 VGDIGD 101

RESULT 16
US-10-224-951-4
; Sequence 4, Application US/10224951
; Publication No. US20030022313A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHN HOPKINS SCHOOL OF MEDICINE
; APPLICANT: WORLEY, PAULI
; APPLICANT: TSUI, CYNTHIA
; TITLE OF INVENTION: NOVEL NEURONAL CELL GROWTH FACTOR
; FILE REFERENCE: JHU 1340-1 (07265/086002)
; CURRENT APPLICATION NUMBER: US/10/224,951
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US/09/098,358
; PRIOR FILING DATE: 1998-06-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: RAT CRP

US-10-224-951-4

Query Match 96.9%; Score 31; DB 14; Length 211;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
Db 151 VGDIGD 156

RESULT 17
US-10-739-930-7500
; Sequence 7500, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 7500
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(276)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C168951_3.p
US-10-739-930-7500

Query Match 96.9%; Score 31; DB 16; Length 276;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
Db 172 VGDIGD 177

RESULT 18
US-10-739-930-7499
; Sequence 7499, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 7499
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(299)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C168951_1.p
US-10-739-930-7499

Query Match 96.9%; Score 31; DB 16; Length 299;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6

Db 144 VGDIGD 149

RESULT 19

US-10-156-761-11418
Sequence 11418, Application US/10156761

Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11418
LENGTH: 328
TYPE: PRT
ORGANISM: Streptomyces avermitilis

US-10-156-761-11418

Query Match
Best Local Similarity 96.9%; Score 31; DB 14; Length 328;
Matches 5; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6

Db 105 VGDIGD 110

RESULT 20

US-10-389-566-367
Sequence 367, Application US/10389566

Publication No. US20040025202A1
GENERAL INFORMATION:
APPLICANT: Monsanto Technology, LLC
APPLICANT: Laurie, Cathy C
TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
FILE REFERENCE: 38-77(52900)D
CURRENT APPLICATION NUMBER: US/10/389,566
CURRENT FILING DATE: 2003-03-31
PRIOR APPLICATION NUMBER: US 60/365,301
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/391,786
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/392,018
PRIOR FILING DATE: 2002-06-26
NUMBER OF SEQ ID NOS: 2459
SOFTWARE: PatentIn version 3.2
SEQ ID NO 367
LENGTH: 342
TYPE: PRT
ORGANISM: Zea mays

NAME/KEY: misc_feature
LOCATION: (1)..(1)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10)..(10)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc_feature

LOCATION: (258)..(258)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (302)..(302)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-367

Query Match
Best Local Similarity 96.9%; Score 31; DB 15; Length 342;
Matches 5; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6

Db 51 VGDIGD 56

RESULT 21

US-10-369-493-23444
Sequence 23444, Application US/10369493

Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 23444
LENGTH: 364
TYPE: PRT
ORGANISM: Deinococcus radiodurans
US-10-369-493-23444

Query Match
Best Local Similarity 96.9%; Score 31; DB 15; Length 364;
Matches 5; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6

Db 70 VGDIGD 75

RESULT 22

US-10-424-599-218355
Sequence 218355, Application US/10424599

Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 218355
LENGTH: 464
TYPE: PRT
ORGANISM: Glycine max

NAME/KEY: unsure
LOCATION: (1)..(464)
OTHER INFORMATION: unsure at all Xaa locations


```
FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39201C.1.pcp
US-10-424-599-218355

Query Match      96.9%; Score 31; DB 15; Length 464;
Best Local Similarity 83.3%; Pred. No. 6.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :|||||
Db      320 VGDIGD 325

RESULT 23
US-10-425-115-254563
; Sequence 254563, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 254563
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_163740C.1.pcp
US-10-425-115-254563

Query Match      96.9%; Score 31; DB 16; Length 537;
Best Local Similarity 83.3%; Pred. No. 7.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :|||||
Db      318 VGDIGD 323

RESULT 24
US-10-369-493-11876
; Sequence 11876, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11876
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-11876

Query Match      96.9%; Score 31; DB 15; Length 693;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 IGDIGD 6
       :|||||
Db      409 IGDIGD 414

RESULT 25
US-10-747-133-1
; Sequence 1, Application US/10747133
; Publication No. US20040229309A1
; GENERAL INFORMATION:
; APPLICANT: Sakaguchi, Nobuo
; APPLICANT: Kuwahara, Kazuhiko
; APPLICANT: Kuwahara, Ganp Protein
; TITLE OF INVENTION: GANP Protein
; FILE REFERENCE: 050208-0014
; CURRENT APPLICATION NUMBER: US/10/747,133
; CURRENT FILING DATE: 2003-12-30
; PRIOR APPLICATION NUMBER: US/09/914,272
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: PCT/JP99/04634
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 47035/1999
; PRIOR FILING DATE: 1999-02-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1971
; TYPE: PRT
; ORGANISM: Mouse
US-10-747-133-1

Query Match      96.9%; Score 31; DB 16; Length 1971;
Best Local Similarity 83.3%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :|||||
Db      1388 VGDIGD 1393

RESULT 26
US-10-424-599-235305
; Sequence 235305, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 235305
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_54508C.1.pcp
US-10-424-599-235305

Query Match      93.8%; Score 30; DB 15; Length 115;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :|||||
Db      38 IGDIGD 43

RESULT 27
US-10-732-923-4243
```

```
; Sequence 4243, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 4243
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Sphaerechinus granularis
; US-10-732-923-4243

Query Match          93.8%; Score 30; DB 17; Length 115;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
        |||:|
Db      44 IGDIGD 49

RESULT 28
US-10-425-115-185817
; Sequence 185817, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 185817
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_101051C.1.pep
; US-10-425-115-185817

Query Match          93.8%; Score 30; DB 16; Length 204;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
        |||:|
Db      132 IGDIGD 137

RESULT 29
US-10-425-115-335683
; Sequence 335683, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
```

```
; SEQ ID NO 335683
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_69249C.1.pep
; US-10-425-115-335683

Query Match          93.8%; Score 30; DB 16; Length 247;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
        :|||:|
Db      202 IGDIGD 207

RESULT 30
US-10-425-114-52030
; Sequence 52030, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52030
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700153786_FLI.pep
; US-10-425-114-52030

Query Match          93.8%; Score 30; DB 15; Length 328;
Best Local Similarity 66.7%; Pred. No. 7.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
        :|||:|
Db      243 VGDVGD 248

RESULT 31
US-10-425-114-68337
; Sequence 68337, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68337
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
```

OTHER INFORMATION: Clone ID: UC-ZMROB73052A10_F11.pdp
US-10-425-114-68337

Query Match
Best Local Similarity 93.8%; Score 30; DB 15; Length 477;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
: : : :
Db 236 VGDVGD 241

RESULT 32

US-10-425-114-57203
Sequence 57203, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 57203
LENGTH: 486
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLB73152G04_F11.pdp
US-10-425-114-57203

Query Match
Best Local Similarity 93.8%; Score 30; DB 15; Length 486;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
: : : :
Db 5 IGDIGD 10

RESULT 33

US-10-425-115-325328
Sequence 325328, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 325328
LENGTH: 486
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(486)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_59770C.1.pdp
US-10-425-115-325328

Query Match
Best Local Similarity 93.8%; Score 30; DB 16; Length 486;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
: : : :
Db 5 IGDIGD 10

RESULT 34

US-10-425-115-239897
Sequence 239897, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 239897
LENGTH: 540
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_150370C.1.pdp
US-10-425-115-239897

Query Match
Best Local Similarity 93.8%; Score 30; DB 16; Length 540;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
: : : :
Db 155 VGDVGD 160

RESULT 35

US-10-424-599-170978
Sequence 170978, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 170978
LENGTH: 567
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(567)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_125408C.1.pdp
US-10-424-599-170978

Query Match
Best Local Similarity 93.8%; Score 30; DB 15; Length 567;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6

Db 87 IGDGD 92

RESULT 36

US-10-437-963-184415
; Sequence 184415, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 184415
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) - (568)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81409C.1.pep
; US-10-437-963-184415

Query Match 93.8%; Score 30; DB 16; Length 568;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDGD 6
: : : :
Db 183 VGDVGD 188

RESULT 37

US-10-437-963-139894
; Sequence 139894, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 139894
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41143C.1.pep
; US-10-437-963-139894

Query Match 93.8%; Score 30; DB 16; Length 598;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IGDGD 6
: : : :
Db 213 VGDVGD 218

RESULT 38

US-10-437-963-181573
; Sequence 181573, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 181573
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78833C.1.pep
; US-10-437-963-181573

Query Match 93.8%; Score 30; DB 16; Length 664;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDGD 6
: : : :
Db 166 VGDVGD 171

RESULT 39

US-10-937-730A-6
; Sequence 6, Application US/10937730A
; Publication No. US20050112726A1
; GENERAL INFORMATION:
; APPLICANT: HU, Zhihao
; APPLICANT: REID, Ralph
; TITLE OF INVENTION: Biosynthetic Gene Cluster for Leptomycin
; FILE REFERENCE: 010092.02
; CURRENT APPLICATION NUMBER: US/10/937,730A
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: 60/502,423
; PRIOR FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: 60/553,384
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3415
; TYPE: PRT
; ORGANISM: Streptomyces sp. ATCC 39366
; US-10-937-730A-6

Query Match 93.8%; Score 30; DB 17; Length 3415;
Best Local Similarity 66.7%; Pred. No. 8.4e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDGD 6
: : : :
Db 1930 VGDVGD 1935

```
RESULT 40
US-10-424-599-227900
; Sequence 227900, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 227900
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_47821C.1.pcp
US-10-424-599-227900

Query Match          90.6%; Score 29; DB 15; Length 73;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :||:|
Db      25 VGDLDG 30

RESULT 41
US-10-425-115-253571
; Sequence 253571, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 253571
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_162838C.1.pcp
US-10-425-115-253571

Query Match          90.6%; Score 29; DB 16; Length 79;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :||:|
Db      31 LGDVGD 36

RESULT 42
US-10-424-599-263936
; Sequence 263936, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 263936
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80353C.1.pcp
US-10-424-599-263936

Query Match          90.6%; Score 29; DB 15; Length 90;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :||:|
Db      76 IGDWGD 81

RESULT 43
US-10-425-115-355214
; Sequence 355214, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 355214
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(152)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_87122C.1.pcp
US-10-425-115-355214

Query Match          90.6%; Score 29; DB 16; Length 152;
Best Local Similarity 66.7%; Pred. No. 5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
       :||:|
Db      58 VGDLDG 63

RESULT 44
US-10-767-701-60220
; Sequence 60220, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
```

```
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 60220
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 7658982.pcp
US-10-767-701-60220
```

```
Query Match          90.6%; Score 29; DB 16; Length 174;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 IGDIGD 6
        :|||:
Db       108 LGDVGD 113
```

```
RESULT 45
US-10-437-963-14654
; Sequence 14654, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 14654
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_47258C.1.pcp
US-10-437-963-14654
```

```
Query Match          90.6%; Score 29; DB 16; Length 187;
Best Local Similarity 66.7%; Pred. No. 6.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 IGDIGD 6
        :|||:
Db       67 LGDVGD 72
```

```
RESULT 46
US-10-425-115-201329
; Sequence 201329, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 363326
; SEQ ID NO 201329
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Oryza sativa
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115198C.1.pcp
US-10-425-115-201329
```

```
Query Match          90.6%; Score 29; DB 16; Length 189;
Best Local Similarity 83.3%; Pred. No. 6.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 IGDIGD 6
        :|||:
Db       156 MEDIGD 161
```

```
RESULT 47
US-10-437-963-156021
; Sequence 156021, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156021
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55729C.1.pcp
US-10-437-963-156021
```

```
Query Match          90.6%; Score 29; DB 16; Length 321;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 IGDIGD 6
        :|||:
Db       279 VGDIGD 284
```

```
RESULT 48
US-10-437-963-102521
; Sequence 102521, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 102521
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Oryza sativa
```

FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100035C.1.pep
US-10-437-963-102521

Query Match 90.6%; Score 29; DB 16; Length 355;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
:|:|
Db 275 VGDIGD 280

RESULT 49
US-10-156-761-14669
; Sequence 14669, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14669
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14669

Query Match 90.6%; Score 29; DB 14; Length 393;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
:|:|
Db 213 LGDVGD 218

RESULT 50
US-10-425-114-45513
; Sequence 45513, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: LIU, JINGDONG
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45513
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700451402_FLI.pep
US-10-425-114-45513

Query Match 90.6%; Score 29; DB 15; Length 422;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
:|:|
Db 246 LGDVGD 251

Search completed: July 27, 2005, 01:24:19
Job time : 154 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 01:01:39 ; Search time 39 seconds
(without alignments)
14.803 Million cell updates/sec

Title: US-10-027-015A-4
Perfect score: 32
Sequence: 1 IGDIGD 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	71	2	D36943
2	32	100.0	190	2	F97008
3	32	100.0	225	2	S16584
4	31	96.9	166	1	B35703
5	31	96.9	166	2	A53812
6	31	96.9	225	2	A40326
7	31	96.9	225	2	A31583
8	31	96.9	225	2	JK0259
9	31	96.9	230	2	A42579
10	31	96.9	284	2	T23528
11	31	96.9	364	2	C75597
12	31	96.9	807	2	D97810
13	31	96.9	807	2	A71663
14	30	93.8	156	2	H71052
15	30	93.8	208	1	F70485
16	30	93.8	208	2	A12674
17	30	93.8	319	1	K1BSFF
18	30	93.8	406	2	H69143
19	30	93.8	478	1	S61986
20	30	93.8	497	2	S60161
21	30	93.8	505	2	S74648
22	30	93.8	547	2	A36046
23	29	90.6	84	2	B86209
24	29	90.6	173	2	A70914
25	29	90.6	198	2	AG3239
26	29	90.6	218	2	H81283
27	29	90.6	318	2	H83610
28	29	90.6	320	2	T36623
29	29	90.6	338	2	T23006

30	29	90.6	347	2	T34919	oxidoreductase - S
31	29	90.6	427	2	AE1305	conserved hypothet
32	29	90.6	441	2	C28379	nitrogenase (EC 1.
33	29	90.6	441	2	F95321	nitrogenase (EC 1.
34	29	90.6	684	2	T36771	probable integral
35	29	90.6	695	2	D84634	hypothetical prote
36	28	87.5	35	2	T06376	ferredoxin-NADP re
37	28	87.5	115	2	D69188	nitrogen regulator
38	28	87.5	147	2	B82523	hypothetical prote
39	28	87.5	154	2	F95938	probable transcrip
40	28	87.5	183	2	D69108	phycocyanin alpha
41	28	87.5	196	2	T15917	hypothetical prote
42	28	87.5	212	2	F84481	Mutator-like trans
43	28	87.5	222	2	H69007	tungsten formylmet
44	28	87.5	231	2	G84722	hypothetical prote
45	28	87.5	256	2	G97778	cell surface antiq
46	28	87.5	285	2	D97094	probable oxidoredu
47	28	87.5	295	2	G81785	hypothetical prote
48	28	87.5	296	2	C82640	hypothetical prote
49	28	87.5	303	2	B84744	hypothetical prote
50	28	87.5	324	2	B83043	probable glycosyl

ALIGNMENTS

RESULT 1

D36943 chemoreceptor protein dcrd - Desulfovibrio vulgaris (strain Hildenborough) (fragment)
C:Species: Desulfovibrio vulgaris
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C:Accession: D36943
R:Decker, H.M., Voordouw, G.
J. Bacteriol. 176, 351-358, 1994
A:Title: Identification of a large family of genes for putative chemoreceptor proteins i
A:Reference number: A36943; MUID:94117369; PMID:828529
A:Accession: D36943
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <DEC>
A:Cross-references: UNIPROT:Q46589; GB:L26013
A>Note: authors translated the codon GGC for residue 9 as Ala, GGT for residue 47 as Thr
TG for residue 62 as Ser, ACC for residue 65 as His, AAC for residue 66 as Ser, GTC for

Query Match	Score	DB 2;	Length
Best Local Similarity	100.0%;	Pred. No. 9;	71;
Matches	6;	Conservative	0;
		Mismatches	0;
		Indels	0;
		Gaps	0;

RESULT 2

F97008 hypothetical protein CAC0881 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: F97008
R:Dalling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: F97008
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <KUR>
A:Cross-references: UNIPROT:Q97KN6; GB:AE001437; PIDN:AAK78857.1; PID:G15023777; GSPDB:C
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0881

Query Match 100.0%; Score 32; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 95 IGDIGD 100

RESULT 3

S16584
coat protein - peanut stunt virus
C/Species: peanut stunt virus
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S16584
R/Naidu, R.A.; Collins, G.B.; Ghahrial, S.A.
Plant Mol. Biol. 17, 175-177, 1991
A/Title: Nucleotide sequence analysis of a cDNA clone encoding the coat protein gene of
A/Reference number: S16584; MUID:91329730; PMID:1868220
A/Accession: S16584
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-225 <NAI>
A/Cross-references: UNIPROT:Q89526; EMBL:X56544; NID:g61371; PIDN:CAA39885.1; PID:g61372
C/Superfamily: cucumber mosaic virus coat protein

Query Match 100.0%; Score 32; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 175 IGDIGD 180

RESULT 4

B35703
cofilin - chicken
C/Species: Gallus gallus (chicken)
C/Date: 12-Oct-1990 #sequence_revision 27-Jun-1994 #text_change 22-Jun-1999
C/Accession: B35703
R/Abe, H.; Endo, T.; Yamamoto, K.; Obinata, T.
Biochemistry 29, 7420-7425, 1990
A/Title: Sequence of cDNAs encoding actin depolymerizing factor and cofilin of embryonic
A/Reference number: A35703; MUID:91027755; PMID:1699599
A/Accession: B35703
A/Molecule type: mRNA
A/Residues: 1-166 <ABE>
A/Cross-references: GB:M55659; NID:g211569; PIDN:AAA62732.1; PID:g211570; GB:U02915
C/Comment: Cofilin reversibly regulates actin polymerization and depolymerization in a F
C/Superfamily: cofilin
C/Keywords: actin binding; phosphoprotein
F/26-36/Region: nuclear location signal
F/104-134/Region: actin binding #status predicted

Query Match 96.9%; Score 31; DB 1; Length 166;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 57 VGDIGD 62

RESULT 5

A53812
cofilin, muscle - mouse
C/Species: Mus musculus (house mouse)
C/Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C/Accession: A53812
R/Ono, S.; Minami, N.; Abe, H.; Obinata, T.
J. Biol. Chem. 269, 15280-15286, 1994

A/Title: Characterization of a novel cofilin isoform that is predominantly expressed in n
A/Reference number: A53812; MUID:94253093; PMID:8195165
A/Accession: A53812
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-166 <ONO>
A/Cross-references: UNIPROT:P45591; GB:I29468; NID:g498016; PIDN:AAA37433.1; PID:g498017
C/Superfamily: cofilin
C/Keywords: actin binding; muscle

Query Match 96.9%; Score 31; DB 2; Length 166;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 57 VGDIGD 62

RESULT 6

A40326
C-reactive protein precursor - golden hamster
C/Species: Mesocricetus auratus (golden hamster)
C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C/Accession: A40326; B40326
R/Dowton, S.B.; Holden, S.N., 1991
Biochemistry 30, 9531-9538, 1991
A/Title: C-reactive protein (CRP) of the Syrian hamster.
A/Reference number: A40326; MUID:91369977; PMID:1892852
A/Accession: A40326
A/Molecule type: DNA
A/Residues: 1-225 <DOM>
A/Cross-references: UNIPROT:P49262; GB:U05343; GB:S56005
A/Note: the authors translated the codon CTA for residue 199 as Asp and GAC for residue ;
A/Accession: B40326
A/Molecule type: mRNA
A/Residues: 1-225 <DOM2>
A/Cross-references: GB:U05343; GB:S56005
C/Comment: This protein, unlike its homolog SAP (serum amyloid P component), is not glyco
C/Genetics:
A/Introns: 22/1
C/Superfamily: C-reactive protein
C/Keywords: pentamer
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-225/Product: C-reactive protein #status predicted <MAT>

Query Match 96.9%; Score 31; DB 2; Length 225;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 172 VGDIGD 177

RESULT 7

A31583
C-reactive protein precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C/Accession: A31583; S06286
R/Onishi, S.; Maeda, S.; Nishiguchi, S.; Arao, T.; Shimada, K.
Biochem. Biophys. Res. Commun. 156, 814-822, 1988
A/Title: Structure of the mouse C-reactive protein gene.
A/Reference number: A31583; MUID:89050112; PMID:3190681
A/Accession: A31583
A/Molecule type: DNA
A/Residues: 1-225 <OHN>
A/Cross-references: UNIPROT:P14847; GB:X13588; NID:g50571; PIDN:CAA31928.1; PID:g295904
R/Mittlehead, A.S.; Zahedi, K.; Rits, M.; Mortensen, R.F.; Lelias, J.M.
Biochem. J. 266, 283-290, 1990
A/Title: Mouse C-reactive protein. Generation of cDNA clones, structural analysis, and ir
A/Reference number: S08286; MUID:90197719; PMID:2310378

A:Accession: S08286
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-133, 'P', 135-225 <WHI>
A:Cross-references: EMBL:X17496; NID:G50563; PIDN:CAA35531.1; PID:G50564
C:Genetics:
A:Introns: 22/1
C:Superfamily: C-reactive protein

Query Match 96.9%; Score 31; DB 2; Length 225;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 172 VGDIGD 177

RESULT 8
C-reactive protein precursor - guinea pig
C:Species: *Cavia porcellus* (guinea pig)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: JX0259
R:Rubio, N.; Sharp, P.M.; Rits, M.; Zahedi, K.; Whitehead, A.S.
J. Biochem. 113, 277-284, 1993
A:Title: Structure, expression, and evolution of guinea pig serum amyloid P component at F120-225/Product: C-reactive protein #status predicted <MET>
F155-116/Disulfide bonds: #status predicted
Query Match 96.9%; Score 31; DB 2; Length 225;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 172 VGDIGD 177

RESULT 9
A42579
C-reactive protein precursor - rat
N:Alternate names: CRP
C:Species: *Rattus norvegicus* (Norway rat)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A42579
R:Rasoulizadeh, M.; Sambasivam, H.; Azadi, P.; Dell, A.; Morris, H.R.; Nagpurkar, A.; Mookerjee, J. Biol. Chem. 267, 2947-2954, 1992
A:Title: Derivation of the amino acid sequence of rat C-reactive protein from cDNA clone
A:Reference number: A42579; MUID:92147634; PMID:137750
A:Accession: A42579
A:Molecule type: mRNA
A:Residues: 1-230 <RAS>
A:Cross-references: UNIPROT:P48199; GB:M83176; NID:G203591; PIDN:AAA40964.1; PID:G203592
A>Note: sequence extracted from NCBI backbone (NCBI:P:80758)
C:Superfamily: C-reactive protein
C:Keywords: acute phase; liver; pentamer; pentraxin; plasma

Query Match 96.9%; Score 31; DB 2; Length 230;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6

Db :|||||
170 VGDIGD 175

RESULT 10
T29528
hypochemical protein F27C1.8 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29528
R:Wu, X.; Le, T.T.
Submitted to the EMBL Data Library, November 1996
A:Description: The sequence of *C. elegans* cosmid F27C1.
A:Reference number: Z20636
A:Accession: T29528
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-284 <WUX>
A:Cross-references: UNIPROT:P91285; EMBL:U80441; PIDN:AAB37656.1; GSPDB:GN00019; CESP:F2
A:Experimental source: strain Bristol N2; clone F27C1
C:Genetics:
A:Gene: CESP:F27C1.8
A:Map position: 1

Query Match 96.9%; Score 31; DB 2; Length 284;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 178 VGDIGD 183

RESULT 11
C75597
thymidine diphosphoglucose 4,6-dehydratase - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: C75597
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utechtack, T.; Zaleski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75597
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-364 <WHI>
A:Cross-references: UNIPROT:Q9R2B3; GB:AE001862; GB:AE001825; NID:G6460468; PIDN:AAF1226
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0041
A:Map position: 2
C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 96.9%; Score 31; DB 2; Length 364;
Best Local Similarity 83.3%; Pred. No. 90;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 70 VGDIGD 75

RESULT 12
D97810
DNA topoisomerase (ATP-hydrolysing) (EC 5.99.1.3) - *Rickettsia conorii* (strain Malish 7)
C:Species: *Rickettsia conorii*
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: D97810
R:Gat, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.B.; Barbe, V.; Samson, D.; Re
Science 293, 2093-2098, 2001

A/Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.
A/Reference number: A97700; MUID:21442074; PMID:11557893
A/Accession: D97810
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-807 <KIR>
A/Cross-references: UNIPROT:Q92H87; GB:AE006914; P1DN:AA03422.1; P1D:G15619989; GSPDB:C
C/Genetics:
A/Gene: gyvB2
C/Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B
C/Keywords: isomerase

Query Match 96.9%; Score 31; DB 2; Length 807;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 35 IGDVGD 40

RESULT 13

A/Title: DNA gyrase chain B (gyvB2) RP580 - *Rickettsia prowazekii*
C/Species: *Rickettsia prowazekii*
C/Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C/Accession: A71663
R/Andersson, S.G.E.; Zomorrodiipour, A.; Andersson, J.O.; Stcheritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A/Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.
A/Reference number: A71630; MUID:99039499; PMID:9823893
A/Accession: A71663
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-807 <AND>
A/Cross-references: UNIPROT:Q92CX2; GB:AJ235272; GB:AJ235269; NID:93861033; P1DN:CAA1502
A/Experimental source: strain Madrid E
C/Genetics:
A/Gene: gyvB2; RP580
C/Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 96.9%; Score 31; DB 2; Length 807;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 35 IGDVGD 40

RESULT 14

A/Title: hypothetical protein PH119 - *Pyrococcus horikoshii*
C/Species: *Pyrococcus horikoshii*
C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C/Accession: H71052
R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohnuki, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A/Reference number: A71000; MUID:98344137; PMID:9679194
A/Accession: H71052
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-156 <KAM>
A/Cross-references: UNIPROT:O58846; GB:AP000005; NID:93236132; P1DN:BAA30218.1; P1D:9325
A/Experimental source: strain OT3
A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
C/Genetics:
A/Gene: PH119

Query Match 93.8%; Score 30; DB 2; Length 156;
Best Local Similarity 83.3%; Pred. No. 55;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IGDIGD 6
Db 30 IGDIGD 35

RESULT 15

A/Title: uracil phosphoribosyltransferase (EC 2.4.2.9) upp-type - *Aquifex aeolicus*
F70485
uracil phosphoribosyltransferase (EC 2.4.2.9) upp-type - *Aquifex aeolicus*
C/Species: *Aquifex aeolicus*
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: F70485
R/Dockert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
V.
Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
A/Reference number: A70300; MUID:98196666; PMID:9537320
A/Accession: F70485
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-208 <AQF>
A/Cross-references: UNIPROT:O67914; GB:AE000776; NID:92984355; P1DN:AA07880.1; P1D:92984
A/Experimental source: strain VF5
C/Genetics:
A/Gene: urap
C/Superfamily: uracil phosphoribosyltransferase upp
C/Keywords: glycosyltransferase; pentosyltransferase

Query Match 93.8%; Score 30; DB 1; Length 208;
Best Local Similarity 83.3%; Pred. No. 76;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 195 IGDIGD 200

RESULT 16

A/Title: hypothetical protein Atu0799 (imported) - *Agrobacterium tumefaciens* (strain C58, Dupont)
A12674
hypothetical protein Atu0799 (imported) - *Agrobacterium tumefaciens* (strain C58, Dupont)
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: A12674
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, B.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: A12674
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-208 <KIR>
A/Cross-references: UNIPROT:Q8UH84; GB:AE008688; P1DN:AA141815.1; P1D:917739171; GSPDB:CF
C/Genetics:
A/Experimental source: strain C58 (Dupont)
A/Gene: Atu0799
A/Map position: circular chromosome

Query Match 93.8%; Score 30; DB 2; Length 208;
Best Local Similarity 83.3%; Pred. No. 76;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 164 IGDIGD 169

RESULT 17

KIBSFF

6-phosphofructokinase (EC 2.7.1.11) - *Bacillus stearothermophilus*
 N:Alternate names: phosphofructokinase 1; phosphohexokinase
 C:Species: *Bacillus stearothermophilus*
 C>Date: 30-Sep-1980 #sequence revision 14-Jul-1994 #text_change 09-Jul-2004
 C:Accession: A27474; A00605; S29784
 R:French, B.A.; Chang, S.H.
 Gene 54, 65-71, 1987
 A:Title: Nucleotide sequence of the phosphofructokinase gene from *Bacillus stearothermophilus*
 A:Reference number: A27474; MUID:87277434; PMID:2956156
 A:Accession: A27474
 A:Molecule type: DNA
 A:Residues: 1-319 <FRE>
 A:Cross-references: UNIPROT:P00512; GB:M15643; NID:g143311; PIDN:AAA22656.1; PID:g143312
 R:Kolb, E.; Hudson, P.J.; Harris, J.I.
 Eur. J. Biochem. 108, 587-597, 1980
 A:Title: Phosphofructokinase: complete amino-acid sequence of the enzyme from *Bacillus stearothermophilus*
 A:Reference number: A91097; MUID:81003881; PMID:6447595
 A:Accession: A00605
 A:Molecule type: protein
 A:Residues: 1-11, N', 13-35, 39-42, V', 44-81, E', 83-94, Q', 96-224, L', 226-263, SA', 266-319
 A>Note: this enzyme is regulated by ADP and phosphoenolpyruvate, in contrast to the eukaryotic enzyme.
 R:Evans, P.R.; Hudson, P.J.
 Nature 279, 500-504, 1979
 A:Title: Structure and control of phosphofructokinase from *Bacillus stearothermophilus*.
 A:Reference number: A93208; MUID:7919719; PMID:156307
 A:Accession: A93208
 A:Contents: annotation; X-ray crystallography, 2.4 angstroms
 A>Note: each chain of the active enzyme consists of two domains; domain 1 consists approximately of the 6-phosphate group of fructose 6-phosphate (F6P) is bound by His-246 from one d to Asp-124; the sugar ring is bound by Arg-249, Met-166, Glu-219, and Asp-124. ATP is bound to the main interaction with ADP appears to be with the beta-phosphate group, which 2 and Thr-155; the phosphate group of phosphoenolpyruvate appears to bind in the same place.
 R:Sakai, H.; Ohta, T.
 Eur. J. Biochem. 211, 851-859, 1993
 A:Title: Molecular cloning and nucleotide sequence of the gene for pyruvate kinase of *Bacillus stearothermophilus*
 A:Reference number: S29784; MUID:93170322; PMID:8436141
 A:Accession: S29784
 A:Molecule type: DNA
 A:Residues: 300-319 <SAK>
 A:Cross-references: EMBL:D13095
 C:Genetics:
 A:Gene: Bg-pfk
 A:Start codon: GTG
 C:Superfamily: ATP-dependent phosphofructokinase, prokaryotic type: 6-phosphofructokinase
 C:Keywords: allosteric regulation; ATP; glycolysis; homotetramer; phosphotransferase
 F:4-278/Domain: 6-phosphofructokinase 1 homology <6pp1>
 Query Match 93.8%; Score 30; DB 1; Length 319;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IGDIGD 6
 :|||
 Db 54 VGDVGD 59
 :|||
 RESULT 18
 H69143
 C:enzyme F420-reducing hydrogenase, beta subunit homolog - *Methanobacterium thermoautotrophicum*
 C:Species: *Methanobacterium thermoautotrophicum*
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: H69143
 R:Smith, D.R.; Doucette-Stamm, L.A.; DeLonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Oltu, D.; Spadefora, R.; Vicare, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Kt, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct
 A:Reference number: A69000; MUID:98037514; PMID:93771463
 A:Accession: H69143
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-406 <MTH>
 A:Cross-references: UNIPROT:Q26441; GB:AE000819; GB:AE000666; NID:g2621336; PIDN:AB8484

A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH341
 C:Superfamily: formate dehydrogenase chain B; ferredoxin 2[4Fe-4S] homology
 Query Match 93.8%; Score 30; DB 2; Length 406;
 Best Local Similarity 66.7%; Pred. No. 1.6e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IGDIGD 6
 :|||
 Db 4 VGDVGD 9
 :|||
 RESULT 19
 S61986
 A:subtilisin-like proteinase (EC 3.4.21.-) YSP3 precursor - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: protein O251; protein UNC478; protein YOR003w; subtilisin-like prote
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: S61986; S66868; S72135
 R:Stekky, F.; Uhlen, M.
 Submitted to the EMBL Data Library, December 1995
 A:Reference number: S61986
 A:Accession: S61986
 A:Molecule type: DNA
 A:Residues: 1-478 <STG>
 A:Cross-references: UNIPROT:P25036; EMBL:U43491; NID:g1150992; PIDN:AA049482.1; PID:g115
 R:Pettersson, B.; Stekky, F.; Uhlen, M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66868
 A:Accession: S66868
 A:Molecule type: DNA
 A:Residues: 1-478 <PST>
 A:Cross-references: EMBL:Z74911; NID:g1420091; PIDN:CAA99191.1; PID:g1420092; GSPDB:GN00
 A:Experimental source: strain S288C
 R:Stekky, F.; Holmberg, A.; Pettersson, B.; Uhlen, M.
 Yeast 12, 1091-1095, 1996
 A:Title: The sequence of a 30 kb fragment on the left arm of chromosome XV from *Saccharo*
 A:Reference number: S72130; MUID:97051599; PMID:8896276
 A:Accession: S72135
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-478 <STW>
 A:Cross-references: EMBL:U43491; NID:g1150992; PIDN:AA049482.1; PID:g1150998
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
 C:Genetics:
 A:Gene: SGD:YSP3; MIPS:YOR003W
 A:Cross-references: SGD:S0005529; MIPS:YOR003W
 A:Map position: 15R
 C:Superfamily: subtilisin, subtilisin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-478/Product: subtilisin-like proteinase YSP3 #status predicted <MAT>
 F:204-421/Domain: subtilisin homology <SBT>
 F:213,245,407/Active site: Asp, His, Ser #status predicted
 Query Match 93.8%; Score 30; DB 1; Length 478;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IGDIGD 6
 :|||
 Db 454 LDDIGD 459
 :|||
 RESULT 20
 S60161
 A:transcription factor relB - African clawed frog
 C:Species: *Xenopus laevis* (African clawed frog)
 C>Date: 15-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
 C:Accession: S60161; S60162
 R:Suzuki, K.; Yamamoto, T.; Inoue, J.

submitted to the EMBL Data Library, July 1995
A/Reference number: S60161
A/Accession: S60161
A/Molecule type: mRNA
A/Residues: 1-497 <SUZ>
A/Cross-references: UNIPROT:P51510; EMBL:D63332; NID:g1109657; PID:d1010251; PID:g110965
R/Suzuki, K.; Yamamoto, T.; Inoue, J.
Nucleic Acids Res. 23, 4664-4669, 1995
A/Title: Molecular cloning of cDNA encoding the Xenopus homolog of mammalian RelB.
A/Reference number: S60162; MUID:96103597; PMID:8524658
A/Accession: S60162
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 78-367 <SUW>
A/Cross-references: EMBL:D63332
C/Genetics:
A/Gene: relb
C/Superfamily: rel transforming protein; rel homology
C/Keywords: DNA binding; transcription regulation
F/78-367/Domain: rel homology <REL>

Query Match 93.8%; Score 30; DB 2; Length 497;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
:|||||
Db 372 IGDIGD 377

RESULT 21
S74648
4-alpha-glucanotransferase malQ - Synechocystis sp. (strain PCC 6803)
N/Alternate names: hypothetical protein sl11676
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S74648
R/Kaneo, T.; Sato, S.; Korani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
sp.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S74648
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-505 <RXN>
A/Cross-references: UNIPROT:P72785; EMBL:D90900; GB:AB001339; NID:g1651768; PIDN:BAAL680
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Genetics:
A/Gene: malQ
C/Superfamily: 4-alpha-glucanotransferase

Query Match 93.8%; Score 30; DB 2; Length 505;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
:|||||
Db 22 IGDIGD 27

RESULT 22
A36046
collagen alpha chain - freshwater sponge (Ephydatia muelleri) (fragment)
C/Species: Ephydatia muelleri
C/Date: 16-Nov-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
C/Accession: A36046
R/Exposito, J.Y.; Garrone, R.
Proc. Natl. Acad. Sci. U.S.A. 87, 6669-6673, 1990
A/Title: Characterization of a fibrillar collagen gene in sponges reveals the early evolu
A/Reference number: A36046; MUID:90370847; PMID:2395869

A/Accession: A36046
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-547 <EXP>
A/Cross-references: UNIPROT:P18856; GB:M34640
C/Genetics: 27/3; 63/3; 82/3; 100/3; 118/3; 190/3; 208/3; 262/3; 304/3; 310/3
F/343-547/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 93.8%; Score 30; DB 2; Length 547;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
:|||||
Db 286 VGDVGD 291

RESULT 23
E86209
protein F22G5.16 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: E86209
R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maltl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: E86209
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-84 <STD>
A/Cross-references: UNIPROT:Q9LNM9; GB:AE005172; NID:g8778566; PIDN:AAF79574.1; GSPDB:GNC
C/Genetics:
A/Gene: F22G5.16
A/Map position: 1

Query Match 90.6%; Score 29; DB 2; Length 84;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IGDIGD 6
:|||||
Db 58 VGDVGD 63

RESULT 24
A70914
hypothetical protein RV0057 - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: A70914
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Suleton, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98255987; PMID:9634220
A/Accession: A70914
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-173 <COL>
A/Cross-references: UNIPROT:P71714; GB:Z80775; GB:AL123456; NID:g3250715; PIDN:CAB02535.1
A/Experimental source: strain H37RV
C/Genetics:

A:Gene: RV0057
C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV0057

Query Match 90.6%; Score 29; DB 2; Length 173;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
DB 132 LGDVGD 137

RESULT 25

AG3239
cell filamentation protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont) F
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AG3239
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, M.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG3239
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-198 <KUR>
A:Cross-references: UNIPROT:Q8U638; GB:AE008690; PIDN:AL46333.1; PID:g17744120; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: fic
A:Genome: plasmid

Query Match 90.6%; Score 29; DB 2; Length 198;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
DB 167 LGDVGD 172

RESULT 26

H81283
helix-turn-helix containing protein Cj1387c [imported] - Campylobacter jejuni (strain NC
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81283
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: AB1250; MUID:20150912; PMID:10688204
A:Accession: H81283
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <PAR>
A:Cross-references: UNIPROT:Q9PMK3; GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAM7381
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1387c

Query Match 90.6%; Score 29; DB 2; Length 218;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
DB 126 IGDWGD 131

RESULT 27

H83610
agmatinase PA0288 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H83610
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: H83610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-318 <STO>
A:Cross-references: UNIPROT:Q916K2; GB:AE004466; GB:AE004091; NID:g9946120; PIDN:AA0367
A:Experimental source: strain PA01
C:Genetics:
A:Gene: speB1; PA0288
C:Superfamily: arginase

Query Match 90.6%; Score 29; DB 2; Length 318;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
DB 85 VGDIGD 90

RESULT 28

T36623
hypothetical protein SCH35.36c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36623
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21610
A:Accession: T36623
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-320 <OLI>
A:Cross-references: UNIPROT:Q9X8X0; EMBL:AL078610; PIDN:CA844391.1; GSPDB:GN00070; SCOE
A:Experimental source: strain A3 (2)
C:Genetics:
A:Gene: SCOE: SCH35.36c

Query Match 90.6%; Score 29; DB 2; Length 320;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
DB 50 LGDVGD 55

RESULT 29

T23006
hypothetical protein F59F3.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23006
R:Kerhaw, J.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z19651
A:Accession: T23006
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-338 <WIL>
A:Cross-references: UNIPROT:Q21039; EMBL:Z68005; PIDN:CAA91992.1; GSPDB:GN00028; CESP:F5

A;Experimental source: clone F59F3
 C;Genetics:
 A;Gene: CESP:F59F3.2
 A;Map position: X
 A;Intons: 17/1; 44/1; 205/1; 228/3; 268/1; 297/2; 317/3

Query Match 90.6%; Score 29; DB 2; Length 338;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :||:|
 Db 170 VGDIGD 175

RESULT 30
 T34919
 oxidoreductase - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C;Accession: T34919
 R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, January 1998
 A;Reference number: Z21558
 A;Accession: T34919
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-347 <OLI>
 A;Cross-references: UNIPROT:O54156; EMBL:AL021409; PIDN:CAA16184.1; GSPDB:GN00070; SCOE
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOE:SC3F7.13
 C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 90.6%; Score 29; DB 2; Length 347;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :||:|
 Db 62 VGDIGD 67

RESULT 31
 AE1305
 conserved hypothetical proteins homolog Imo1845 [imported] - listeria monocytogenes (str
 C;Species: listeria monocytogenes
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C;Accession: AE1305
 R;Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fehli, H.
 Science 294, 849-852, 2001
 D.; Jones, L.M.; Karst, U.
 A;Authors: Krefz, U.; Kuhn, M.; Kunz, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schluter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
 A;Title: Comparative genomics of listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AE1305
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-427 <GLA>
 A;Cross-references: UNIPROT:Q8Y655; GB:NC_003210; PIDN:CAC99923.1; PID:GL6411299; GSPDB:
 A;Experimental source: strain EGD-e
 C;Genetics:
 A;Gene: Imo1845
 C;Superfamily: conserved hypothetical protein HI0125

Query Match 90.6%; Score 29; DB 2; Length 427;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :||:|

Db 167 VGDIGD 172

RESULT 32
 C28379
 nitrogenase (EC 1.18.6.1) molybdenum-iron protein beta chain - Rhizobium meliloti
 C;Species: Rhizobium meliloti
 C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
 C;Accession: C28379
 R;Aguliar, O.M.; Reilaender, H.; Arnold, W.; Puhler, A.
 J. Bacteriol. 169, 5393-5400, 1987
 A;Title: Rhizobium meliloti nifN (fixP) gene is part of an operon regulated by a nifA-dei
 A;Reference number: A91851; MUID:86058743; PMID:3316182
 A;Accession: C28379
 A;Molecule type: DNA
 A;Residues: 1-441 <AGU>
 A;Cross-references: UNIPROT:P12781; GB:ML8272; NID:GL52217; PIDN:AAA89520.1; PID:GL52219
 C;Genetics:
 A;Gene: fixP
 A;Superfamily: dinitrogenase beta chain
 C;Keywords: ATP; nitrogen fixation; oxidoreductase

Query Match 90.6%; Score 29; DB 2; Length 441;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :||:|
 Db 361 VGDIGD 366

RESULT 33
 F95321
 nitrogenase (EC 1.18.6.1) NifN - Sinorhizobium meliloti (strain 1021) megaplasmid pSymA
 C;Species: Sinorhizobium meliloti
 C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C;Accession: F95321
 R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.,
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9889, 2001
 A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A;Reference number: A95262; MUID:21396509; PMID:11481432
 A;Accession: F95321
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-441 <KUR>
 A;Cross-references: UNIPROT:P12781; GB:AE006469; PIDN:AAK65136.1; PID:GL4523576; GSPDB:G
 A;Experimental source: strain 1021, megaplasmid pSymA
 R;Gallibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A;Reference number: A96039; MUID:21368234; PMID:11474104
 A;Contents: annotation
 C;Genetics:
 A;Gene: nifN
 A;Genome: plasmid
 C;Superfamily: dinitrogenase beta chain
 C;Keywords: ATP; nitrogen fixation; oxidoreductase

Query Match 90.6%; Score 29; DB 2; Length 441;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :||:|
 Db 361 VGDIGD 366

RESULT 34

T36771
probable integral membrane protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36771
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21574
A:Accession: T36771
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-684 <SEER>
A:Cross-references: UNIPROT:Q9S282; EMBL:AL096844; PIDD:CA550875.1; GSPDB:GN00070; SCOPED
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SC08DB:SC128.01

Query Match
Best Local Similarity 66.7%; Score 29; DB 2; Length 684;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
:|||||
DB 192 LGDVGD 197

RESULT 35
D84634
hypothetical protein At2g24250 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84634
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:1061197
A:Accession: D84634
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-695 <STO>
A:Cross-references: UNIPROT:Q9ZUH0; GB:AE002093; NID:g4115385; PIDD:AAD03386.1; GSPDB:GN
C:Genetics:
A:Gene: At2g24250
A:Map position: 2

Query Match
Best Local Similarity 66.7%; Score 29; DB 2; Length 695;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
:|||||
DB 263 VGDLDG 268

RESULT 36
T06376
ferredoxin-NADP reductase (EC 1.18.1.2) - garden pea (fragment)
C:Species: Pisum sativum (garden pea)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06376
R:Maraffini, L.A.; Ceccarelli, E.A.
submitted to the EMBL Data Library, April 1998
A:Description: Involvement of tyrosine-114 of pea ferredoxin-NADP+ reductase on FAD bind
A:Reference number: Z15636
A:Accession: T06376
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-35 <MAR>
A:Cross-references: UNIPROT:O65208; EMBL:AF057182; NID:g3088653; PIDD:AA014746.1; PIDD:93
C:Genetics:
A:Genome: nuclear

C:Superfamily: ferredoxin-NADP+ reductase; cytochrome-b5 reductase homology
C:Keywords: chloroplast; electron transfer; FAD; flavoprotein; NADP; oxidoreductase; pho

Query Match
Best Local Similarity 87.5%; Score 28; DB 2; Length 35;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
:|||||
DB 3 IGDIGD 8

RESULT 37
D69188
nitrogen regulatory protein P-II MTH664 - Methanobacterium thermoautotrophicum (strain D
C:Species: Methanobacterium thermoautotrophicum
A:Variety: strain Delta H
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D69188
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: D69188
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-115 <MTH>
A:Cross-references: UNIPROT:O26760; GB:AE000846; GB:AE000666; NID:g2621740; PIDD:AAB8516
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH664
A:Start codon: GTG
C:Superfamily: regulatory protein P-II
C:Keywords: phosphoprotein; signal transduction
F.54/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match
Best Local Similarity 87.5%; Score 28; DB 2; Length 115;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IGDIGD 6
:|||||
DB 87 GDIGD 91

RESULT 38
E82523
hypothetical protein XR2732 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82523
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82523
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-147 <SIM>
A:Cross-references: UNIPROT:Q9P918; GB:AE004079; GB:AE003849; NID:g9107960; PIDD:AAF8551
A:Experimental source: strain 9a5c
R:Simson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.; A
Briener, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XP2732

Query Match 87.5%; Score 28; DB 2; Length 147;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDIGD 6
Db 101 LGDGLD 106

RESULT 39

P95938
probable transcription regulator protein [imported] - Sinorhizobium meliloti (strain 102
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: P95938

R;Rinan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernar
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: P95938

A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-154 <KIR>
A;Cross-references: UNIPROT:Q92VD3; GB:A1591985; PIDN:CAQ9174.1; PID:G15140659; GSPDB:C
A;Experimental source: strain 1021, megaplasmid pSymb
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMD21270

A;Genome: plasmid

Query Match 87.5%; Score 28; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
Db 46 GDIGD 50

RESULT 40

D69108
phycoerythrin alpha phycoerythrin lyase CpeC - Methanobacterium thermoautotrophicum (str
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: D69108

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vitcare, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Neelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: D69108

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-183 <MTH>

A;Cross-references: UNIPROT:Q27834; GB:AE000934; GB:AE000666; NID:92622924; PIDN:AAB8627
A;Experimental source: strain Delta H

C;Genetics:
A;Gene: MTH1806

Query Match 87.5%; Score 28; DB 2; Length 183;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDIGD 6
Db 145 LGDGLD 150

RESULT 41

T15917
hypothetical protein EEBD8.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15917

R;Chisoe, S.
submitted to the EMBL Data Library, July 1995
A;Description: The sequence of C. elegans cosmid EEBD8.
A;Reference number: 218428
A;Accession: T15917

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-196 <CHI>

A;Cross-references: UNIPROT:Q09511; EMBL:U23484; NID:9733597; PID:9733604; PIDN:AAC46767
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:EEBD8.7

A;Intons: 69/2; 115/2; 133/3
A;Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
F;20-87/Domain: ribonucleoprotein repeat homology <RMR>

Query Match 87.5%; Score 28; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
Db 43 GDIGD 47

RESULT 42

F84481
Mutator-like transposase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84481

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bent, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84481

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <STO>
A;Cross-references: UNIPROT:Q92V08; GB:AE002093; NID:93779027; PIDN:AAC67206.1; GSPDB:GNC

C;Genetics:
A;Gene: AT2G07030
A;Map position: 2

Query Match 87.5%; Score 28; DB 2; Length 212;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDIGD 6
Db 47 GDIGD 51

RESULT 43
H69007
tungensten formylmethanofuran dehydrogenase, subunit C homolog - Methanobacterium thermoautotrophicum
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 12-Jul-2004
C:Accession: H69007
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Kt, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: H69007
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-222 <MTH>
A:Cross-references: UNIPROT:O26209; GB:AE000800; GB:AE000666; NID:92621130; PIDN:AA8460
C:Genetics:
A:Experimental source: strain Delta H
C:Gene: MTH106
C:Superfamily: glutamate synthase, domain3

Query Match
Best Local Similarity 87.5%; Score 28; DB 2; Length 222;
100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GDIGD 6
|||||
DB 134 GDIGD 138

RESULT 44
G84722
hypochemical protein At2g31600 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: G84722
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Wolf, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84722
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <STO>
A:Cross-references: UNIPROT:Q9SIP8; GB:AE002093; NID:94582454; PIDN:AMD24838.1; GSPDB:GN
C:Genetics:
A:Gene: At2g31600
A:Map position: 2
C:Superfamily: Arabidopsis thaliana hypothetical protein F5K20.160

Query Match
Best Local Similarity 87.5%; Score 28; DB 2; Length 231;
100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GDIGD 6
|||||
DB 142 GDIGD 146

RESULT 45
G97778
cell surface antigen homolog [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97778
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.B.; Barbe, V.; Samson, D.; Ro Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: G97778

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-256 <KUR>
A:Cross-references: UNIPROT:Q92HY9; GB:AE006914; PIDN:AA03169.1; PID:915619717; GSPDB:G
C:Genetics:
A:Gene: RC0631

Query Match
Best Local Similarity 87.5%; Score 28; DB 2; Length 256;
100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GDIGD 6
|||||
DB 230 GDIGD 234

RESULT 46
D97094
probable oxidoreductase, ortholog of GSP39 B. subtilis CAC1576 [imported] - Clostridium
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: D97094
R:Noelling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4833-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A86900; MUID:21359325; PMID:21359325
A:Accession: D97094
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-285 <KUR>
A:Cross-references: UNIPROT:Q97IR4; GB:AE001437; PIDN:AAK79543.1; PID:915024529; GSPDB:G
C:Genetics:
A:Gene: CAC1576
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match
Best Local Similarity 87.5%; Score 28; DB 2; Length 285;
100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GDIGD 6
|||||
DB 99 GDIGD 103

RESULT 47
G81785
hypothetical protein NMA2138 [imported] - Neisseria meningitidis (strain Z2491 serogrou
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: G81785
R:Parhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:2022556; PMID:10761919
A:Accession: G81785
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <PAR>
A:Cross-references: UNIPROT:Q9JSR9; GB:AL162758; GB:AL157959; NID:97380672; PIDN:CAB8535
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA2138

Query Match
Best Local Similarity 87.5%; Score 28; DB 2; Length 295;
100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GDIGD 6
|||||
DB 162 GDIGD 166

RESULT 48

C82640
hypothetical protein XPI170 [imported] - *Xylella fastidiosa* (strain 9a5c)
C/Species: *Xylella fastidiosa*
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: C82640
R/Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: C82640
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-296 <STM>
A/Cross-references: UNIPROT:Q9PCK6; GB:AE003999; GB:AE003849; NID:g9106837; PIDN:AAFB457
A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Aencio, M.; Alvarenga, R.; A
Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; From
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, J.P.; Klieger, J.E.; Kuramae, E.B.; Laig
chado, M.A.; Madalena, A.M.B.N.; Madalena, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J.; de M.; da Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuchiko, M.H.; Valleda, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A/Reference number: A59328
A/Contents: annotation
C/Genetics:
A/Gene: XPI170

Query Match 87.5%; Score 28; DB 2; Length 296;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
DB 189 LGDLGD 194

RESULT 49

B84744
hypothetical protein At2g33330 [imported] - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: B84744
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, P.; Vanhaken, S.E.; Umayam, L.; Tallon, L.
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: B84744
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-303 <STO>
A/Cross-references: UNIPROT:O22784; GB:AE002093; NID:g2459416; PIDN:AA80651.1; GSPDB:GN
C/Genetics:
A/Gene: At2g33330
A/Map position: 2

Query Match 87.5%; Score 28; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GDIGD 6
DB 198 GDIGD 202

RESULT 50

E83043
probable glycosyl transferase PA4819 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C/Species: *Pseudomonas aeruginosa*
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: E83043
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Loiy, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: E83043
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-324 <STO>
A/Cross-references: UNIPROT:Q9HUZ1; GB:AE004895; GB:AE004091; NID:g9951083; PIDN:AA0820
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA4819
C/Superfamily: stress response protein csbB

Query Match 87.5%; Score 28; DB 2; Length 324;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
DB 41 LGDLGD 46

Search completed: July 27, 2005, 01:21:13
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 00:58:59 ; Search time 174 Seconds
(without alignments)
17.658 Million cell updates/sec

Title: US-10-027-015A-4
Perfect score: 32
Sequence: 1 IGDIGD 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Uniprot 03:*
1: uniprot_prot:*
2: uniprot_crem1:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	73	2	046589
2	32	100.0	190	2	097KN6
3	32	100.0	225	2	089526
4	32	100.0	229	2	06C567
5	32	100.0	372	2	0745P8
6	32	100.0	436	2	09KYG6
7	32	100.0	443	2	08G7E9
8	32	100.0	515	2	09HKD4
9	32	100.0	585	2	09FDPV9
10	32	100.0	1380	2	087MP0
11	32	100.0	2186	2	08NM19
12	32	100.0	2186	2	06G8H1
13	31	96.9	166	1	0022_HUMAN
14	31	96.9	166	1	0022_MOUSE
15	31	96.9	166	1	0022_CHICK
16	31	96.9	166	2	06NZM3
17	31	96.9	180	2	085325
18	31	96.9	189	2	09US43
19	31	96.9	201	2	084BH1
20	31	96.9	214	2	09YBP4
21	31	96.9	222	2	019062
22	31	96.9	225	1	CRP_CAVPO
23	31	96.9	225	1	CRP_MESAU
24	31	96.9	225	1	CRP_MOUSE
25	31	96.9	225	2	091XB3
26	31	96.9	230	2	CRP_RAT
27	31	96.9	247	2	08GHM0
28	31	96.9	252	2	084176
29	31	96.9	281	2	041345
30	31	96.9	281	2	042407
31	31	96.9	284	1	CCDS_CABEL

32	31	96.9	308	2	08DGV6	08dgv6	synchococc
33	31	96.9	328	2	067KV0	067kv0	syndiobacte
34	31	96.9	328	2	082GL3	082gl3	streptomyce
35	31	96.9	348	2	09F2Z0	09f2z0	streptomyce
36	31	96.9	364	2	09R2B3	09r2b3	deinococcus
37	31	96.9	588	2	07XTS7	07xts7	oryza sativ
38	31	96.9	693	2	098KR2	098kr2	rhizobium I
39	31	96.9	807	1	GYRB_RICCN	092h87	rickettsia
40	31	96.9	807	1	GYRB_RICPR	092cx2	rickettsia
41	31	96.9	807	2	07P921	07p921	rickettsia
42	31	96.9	807	2	068WP6	068wf6	rickettsia
43	31	96.9	1015	2	07TMA9	07tma9	rattus norv
44	31	96.9	1224	2	08C777	08c777	mus musculu
45	31	96.9	1637	2	07USG5	07usg5	rhodospirill
46	31	96.9	1732	2	06NMV4	06nmv4	mus musculu
47	31	96.9	1971	1	MC3A_MOUSE	09wu09	mus musculu
48	31	96.9	1971	2	07TS87	07ts87	mus musculu
49	31	96.9	1992	2	080TK5	080tk5	mus musculu
50	31	96.9	2251	2	07RNM7	07rnm7	plasmodium

ALIGNMENTS

```
RESULT 1
ID 046589 PRELIMINARY; PRT; 73 AA.
AC 046589;
DT 01-NOV-1996 (TRENBLER. 01, Created)
DT 01-NOV-1996 (TRENBLER. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLER. 25, Last annotation update)
DE Chemoreceptor (Fragment).
GN Name=cdcd;
OS Desulfovibrio vulgaris.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBT_Taxid=881;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hildenborough;
RX MEDLINE=94117369; PubMed=8288529;
RA Deckers H.M., Voordouw G.;
RT "Identification of a large family of genes for putative chemoreceptor
   proteins in an ordered library of the Desulfovibrio vulgaris
   Hildenborough genome."
RL EMBL; L26013; AAA23358.1; -.
DR PIR; D36943; D36943.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR004090; Me_chemotaxis.
DR Pfam; PF00015; MCPs; signal; 1.
DR PRINTS; PR00260; CHEMTRNSDUC.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
FT NON TER 1 73
FT NON TER 73 73
SQ
Query Match 73 AA; 7460 MW; 595542C07F328477 CRC64;
Best local Similarity 100.0%; Score 32; DB 2; Length 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

ID 097KX6 PRELIMINARY; PRT; 190 AA.
AC 097KX6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein CAC0881.
GN OrderedLocustNames=CAC0881.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=1466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng O.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatunov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.,
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007603; AKT78857.1; -.
DR PIR: F97008; F97008.
DR InterPro: IPR009061; Putativ_DNA_bind.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 190 AA; 22079 MW; 7D6D3F4A38C36F10 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 95 IGDIGD 100

RESULT 3
ID 089526 PRELIMINARY; PRT; 225 AA.
AC 089526;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Coat protein.
GN Name-coat protein;
OS Peanut stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=12313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ER;
RA Naidu R.A., Hu C.-C., Pennington R.E., Ghabrial S.A.,
RT "Differentiation of eastern and western strains of peanut stunt
cucumovirus based on satellite RNA support and nucleotide sequence
homology.";
RL Phytopathology 85:502-507(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ER;
RX MEDLINE=91329730; PubMed=1868220;
RA Naidu R.A., Collins G.B., Ghabrial S.A.,
RT "Nucleotide sequence analysis of a cDNA clone encoding the coat
protein gene of peanut stunt virus.";
RL Plant Mol. Biol. 17:175-177(1991).
DR EMBL: U15730; AAA70087.1; -.
DR EMBL: X55544; CAA39885.1; -.
DR PIR: S16584; S16584.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR00247; Cucumovirus_coat.

DR Pfam: PF00760; Cucumo_coat; 1.
DR PRINTS: PR00222; CUCUMOCOAT.
DR ProDom: PD001284; Cucumovirus_coat; 1.
KW Coat protein.
SQ SEQUENCE 225 AA; 25061 MW; 11BC2174FED98BA1 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 175 IGDIGD 180

RESULT 4
ID 06C567 PRELIMINARY; PRT; 229 AA.
AC 06C567;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to tr|Q9C2Y5 Aspergillus awamori Uracll
DE phosphoribosyltransferase.
GN ORFNames=YAL10E20525g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marc C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barney S., Blanchin S., Beckerich J.M., Bayne E., Bleykasten C.,
RA Boissarie A., Boyer J., Catolico L., Confanieret F., de Daruvar A.,
RA Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Gropi A.,
RA Hantlaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikoleki M., Ozas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenilou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.,
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: CR82131; CAG79790.1; -.
DR GO: GO:0004743; F:pyruvate kinase activity; IEA.
DR GO: GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR GO: GO:0009116; P:nucleoside metabolism; IEA.
DR InterPro: IPR000836; PRTtransferase.
DR InterPro: IPR001697; Pyruvate_kinase.
DR Pfam: PF00156; Pribosyltran; 1.
DR ProDom: PD001009; Pyruvate_kinase; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 229 AA; 25083 MW; 3C0F31AD3D8DCB5 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 214 IGDIGD 219

```

RESULT 5
Q745P8      PRELIMINARY;      PRT;      372 AA.
AC 0745P8;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MAP0246c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L.; Bannantine J., Zhang O., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017227; AAS02563.1; -.
DR InterPro; IPR005142; eRF1_3.
DR Pfam; PF03465; eRF1_3; 1.
KM Complete proteome.
SQ SEQUENCE 372 AA; 39863 MW; 0153C98FDBA2A1D4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 2; Length 372;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 292 IGDIGD 297

RESULT 6
Q9KYG6      PRELIMINARY;      PRT;      436 AA.
AC 09KYG6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Putative transferase.
GN ORFNames=SCC61A.27;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Latke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinovich E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417.141-147(2002).
DR EMBL; AL939113; CAB92272.1; -.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0005058; P:biosynthesis; IEA.
KM Complete proteome; Transferase.
SQ SEQUENCE 436 AA; 45551 MW; 7A54166F63838854 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 2; Length 436;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 IGDIGD 6
Db 410 IGDIGD 415

RESULT 7
Q8G7E9      PRELIMINARY;      PRT;      443 AA.
AC 08G7E9;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Histidine kinase sensor of two-component system.
GN OrderedLocustNames=BL0317;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinomycetales; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Scheil M.A., Karmilantzon M., Snel B., Vilanova D., Berger B.,
RA Peesl G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
RA Fridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99.14422-14427(2002).
CC -1 SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; AB014650; AAN24157.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR005467; His_Kinase.
DR InterPro; IPR003661; His_KinA_N.
DR InterPro; IPR009082; His_kin_homodim.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HSKA_1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HSKA_1.
DR PROSITE; PS50109; HIS_KIN; 1.
KM Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KM Transferase.
SQ SEQUENCE 443 AA; 47995 MW; D0B0999FEB73D3E1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 2; Length 443;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 155 IGDIGD 160

RESULT 8
Q9HKD4      PRELIMINARY;      PRT;      515 AA.
AC 09HKD4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DR Acyl-CoA synthetase related protein.
GN OrderedLocustNames=Ta0667;
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmales;
OC Thermoplasmales; Thermoplasma.
OX NCBI_TaxID=2303;

```

RA N
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001; DOI=10.1038/35035069;
RA Repp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RW Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermophilic scavenger Thermoplasma
acidophilum.";
RL Nature 407:508-513(2000).
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
family.
DR EMBL: A145065; CAC11805.1; -
DR GO: GO:0003824; P: catalytic activity; IEA.
DR GO: GO:0008152; P: metabolism; IEA.
DR InterPro: IPR000873; AMP-Bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PRINTS: PR00154; AMPBINDING.
DR PROSITE: PS00455; AMP_BINDING; 1.
KW Complete proteome.
SQ SEQUENCE 515 AA; 57543 MW; ACE8BB89AE423A0B CRC64;

Query Match 100.0%; Score 32; DB 2; Length 515;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 423 IGDIGD 428

RESULT 9
Q9FDV9 PRELIMINARY; PRT; 585 AA.
ID Q9FDV9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE 4-alpha-glucanotransferase.
GN Name=STAI1;
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22631581; PubMed=12746519; DOI=10.1104/pp.102.016527;
RA Mathebled F., Rai J.P., Dauvillée D., Myers A.M., James M.G.,
RA Schlichting R., Giersch C., Ball S.G., D'Hulst C.,
RT "STAI1, a Chlamydomonas reinhardtii locus required for normal starch
granule biogenesis, encodes a disproportionation enzyme. Further
RT evidence for a function of alpha-1,4 glucanotransferases during starch
RT granule biosynthesis in green algae.";
RL Plant Physiol. 132:137-145(2003).
DR EMBL: AF307843; AAG29840.1; -
DR EMBL: AF307842; AAG29839.1; -
DR HESP; O81172; IESW.
DR GO: GO:0004134; F: 4-alpha-glucanotransferase activity; IEA.
DR GO: GO:0016740; F: transferase activity; IEA.
DR GO: GO:0005975; P: carbohydrate metabolism; IEA.
DR InterPro: IPR003385; Glyco_hydro_77.
DR Pfam: PF02446; 4a-glucanotrans; 1.
DR TIGRFAMs: TIGR00217; malQ; 1.
KW Transferrase.
SQ SEQUENCE 585 AA; 65825 MW; PEF89A1EF658E325 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 101 IGDIGD 106

RESULT 10
Q87MP0 PRELIMINARY; PRT; 1380 AA.
ID Q87MP0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein VP2191.
GN OrderedLocustNames=VP2191;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:H6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL: AP005080; BAC60454.1; -
DR InterPro: IPR008941; TPR-like.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1380 AA; 152808 MW; 52197DEF11AF4E7C CRC64;

Query Match 100.0%; Score 32; DB 2; Length 1380;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 1318 IGDIGD 1323

RESULT 11
Q8NM19 PRELIMINARY; PRT; 2186 AA.
ID Q8NM19;
AC Q8NM19;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MW1699 protein.
GN OrderedLocustNames=MW1699;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL: AP004628; BAB95564.1; -
DR GO: GO:0009986; C: cell surface; IEA.
DR GO: GO:0016020; C: membrane; IEA.
DR InterPro: IPR011439; DUF1542.
DR InterPro: IPR011490; FIVAR.
DR InterPro: IPR005877; Gpos_YSTRK.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF07564; DUF1542; 17.
DR Pfam: PF07554; FIVAR; 1.
DR TIGRFAMs: TIGR01167; LPTXG_anchor; 1.
DR TIGRFAMs: TIGR01166; YSTRK_signal; 1.
KW Complete proteome.


```

SQ SEQUENCE 2186 AA; 237934 MW; 621D7875D1D3FB35 CRC64;
Query Match 100.0%; Score 32; DB 2; Length 2186;
Best Local Similarity 100.0%; Pred. NO. 2.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IGDIGD 6
Db 702 IGDIGD 707

RESULT 12
O6G8H1 PRELIMINARY; PRT; 2186 AA.
ID O6G8H1;
AC O6G8H1;
DT 05-JUL-2004 (TREMblrel. 27, Created)
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)
DE Putative surface anchored protein.
GN OrderedLocustName=SAS1682;
OS Staphylococcus aureus (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Feil E.U., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Alkin R., Barton A.,
RA Baon N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corbin C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagsels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrett B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
RU EMBL; BX571857; CAG3485.1; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR InterPro; IPR011439; DUF1542.
DR InterPro; IPR011490; FIVAR.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF07554; DUF1542; 17.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.
RW Complete proteome.
SQ SEQUENCE 2186 AA; 238063 MW; 3A041AF93D45932C CRC64;
Query Match 100.0%; Score 32; DB 2; Length 2186;
Best Local Similarity 100.0%; Pred. NO. 2.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IGDIGD 6
Db 702 IGDIGD 707

RESULT 13
COF2 HUMAN STANDARD; PRT; 166 AA.
ID COF2_HUMAN;
AC O9Y281;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cofilin, muscle isoform (Cofilin-2).
GN Name=CFI2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Jin J., Li G., Hu S., Li W., Yuan J., Qiang B.;
RT "Isolation of two isoforms of human cofilin cDNA.";
RL Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21315834; PubMed=11422377;
RA Thirion C., Stucka R., Mendel B., Grubler A., Jakesch M., Nowak K.J.,
RA Binz N., Lating N.G., Lochmuller H.;
RT "Characterization of human muscle type cofilin (CFI2) in normal and
RT regenerating muscle.";
RL Eur. J. Biochem. 268:3473-3482(2001).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow, Placenta, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mulhany S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heltan E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U.A., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RU EMBL; AF134802; AAD31280.1; -.
DR EMBL; AF134803; AAD31281.1; -.
DR EMBL; AF283513; AAF97934.1; -.
DR EMBL; AF242299; AAF64498.1; -.
-1- FUNCTION: Controls reversibly actin polymerization and
-1- depolymerization in a pH-sensitive manner. It has the ability to
-1- bind G- and F-actin in a 1:1 ratio of cofilin to actin. It is the
-1- major component of intranuclear and cytoplasmic actin rods (By
-1- similarity).
-1- SUBCELLULAR LOCATION: Intranuclear and cytoplasmic (By
-1- similarity).
-1- ALTERNATIVE PRODUCTS:
-1- Event=Alternative splicing; Named isoforms=2;
-1- Comment=Isoforms are identical at the level of the protein
-1- sequence;
-1- Name=CFI2b;
-1- IsoId=Q9Y281-1; Sequence=Displayed;
-1- Name=CFI2a;
-1- IsoId=Q9Y281-2; Sequence=Not described;
-1- TISSUE SPECIFICITY: Isoform CFI2b is expressed predominantly in
-1- skeletal muscle and heart, while isoform CFI2a is expressed in
-1- various tissues.
-1- PTM: The phosphorylation of Ser-24 may prevent recognition of the
-1- nuclear localization signal.
-1- SIMILARITY: Belongs to the actin-binding proteins ADP family.
-1- This SWISS-PROT entry is copyright. It is produced through a collaboration
-1- between the Swiss Institute of Bioinformatics and the EMBL outstation -
-1- the European Bioinformatics Institute. There are no restrictions on its
-1- use by non-profit institutions as long as its content is in no way
-1- modified and this statement is not removed. Usage by and for commercial
-1- entities requires a license agreement (See http://www.isb-sib.ch/announce/
-1- or send an email to license@sib-sib.ch).
-1- EMBL; AF134802; AAD31280.1; -.
-1- EMBL; AF134803; AAD31281.1; -.
-1- EMBL; AF283513; AAF97934.1; -.
-1- EMBL; AF242299; AAF64498.1; -.

```

```
DR EMBL, BC011444; AAH11444.1; -.
DR EMBL, BC022364; AAH22364.1; -.
DR EMBL, BC022876; AAH22876.1; -.
DR HSSP, P18282; 1AK6.
DR Genew, HGNC:1875; CFL2.
DR MIM, 601443; -.
DR InterPro: IPR002108; Actbind_cofin.
DR Pfam, PF00241; Cofilin_ADF, 1.
DR PRINTS; PR00006; COFILIN.
DR ProDom; PD002129; Actbind_cofin; 1.
DR SMART; SM00102; ADF, 1.
DR PROSITE; PS00325; ACTIN DEPOLYMERIZING; 1.
KM Actin-binding; Alternative splicing; Cytoskeleton; Nuclear protein;
KW Phosphorylation.
FT DOMAIN 30 34 Nuclear localization signal (potential).
FT MOD_RES 106 125 Actin-binding (potential).
FT MOD_RES 24 24 Phosphoserine (by similarity).
SQ SEQUENCE 166 AA; 18736 MW; 48B6DCCAE9FE1CC CRC64;

Query Match 96.3%; Score 31; DB 1; Length 166;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 57 VGDIGD 62

RESULT 14
COF2_MOUSE STANDARD; PRT; 166 AA.
AC P45351;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cofilin, muscle isoform (Cofilin-2).
CN Name=Cfl2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H; TISSUE=Skeletal muscle;
RX MEDLINE=94253093; Pubmed=8195165;
RA Ono S., Minami N., Abe H., Ohtsuka T.;
RT "Characterization of a novel cofilin isoform that is predominantly
RT expressed in mammalian skeletal muscle.";
RL J. Biol. Chem. 269:15280-15286(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA Kraussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Tohiyuki S., Carninci P., Mulhaly S.J.,
RA Bork S.A., McKean P.J., McKenna K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Halton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-1- FUNCTION: Controls reversibly actin polymerization and
```

```
CC depolymerization in a pH-sensitive manner. It has the ability to
CC bind G- and F-actin in a 1:1 ratio of cofilin to actin. It is the
CC major component of intranuclear and cytoplasmic actin rods.
CC -1- SUBCELLULAR LOCATION: Intranuclear and cytoplasmic.
CC -1- TISSUE SPECIFICITY: Predominantly expressed in skeletal muscle.
CC -1- PTM: The phosphorylation of Ser-24 may prevent recognition of the
CC nuclear localization signal.
CC -1- SIMILARITY: Belongs to the actin-binding proteins ADF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L29468; AAA37433.1; -.
DR EMBL, BC007138; AAH07138.1; -.
DR PIR, A53812; A53812.
DR HSSP, P18282; 1AK6.
DR MGD; MGI:101763; CFL2.
DR InterPro: IPR002108; Actbind_cofin.
DR Pfam; PF00241; Cofilin_ADF, 1.
DR PRINTS; PR00006; COFILIN.
DR ProDom; PD002129; Actbind_cofin; 1.
DR SMART; SM00102; ADF, 1.
DR PROSITE; PS00325; ACTIN DEPOLYMERIZING; 1.
KM Actin-binding; Cytoskeleton; Nuclear protein; Phosphorylation.
FT DOMAIN 30 34 Nuclear localization signal (potential).
FT MOD_RES 106 125 Actin-binding (potential).
FT MOD_RES 24 24 Phosphoserine (by similarity).
SQ SEQUENCE 166 AA; 18709 MW; 48B6D7E5AE9FE1CC CRC64;

Query Match 96.3%; Score 31; DB 1; Length 166;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 57 VGDIGD 62

RESULT 15
COF1_CHICK STANDARD; PRT; 166 AA.
AC P21566;
DT 01-MAY-1991 (Rel. 18, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cofilin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=91027755; Pubmed=1699599;
RA Abe H., Endo T., Yamamoto K., Ohtsuka T.;
RT "Sequence of cDNAs encoding actin depolymerizing factor and cofilin of
RT embryonic chicken skeletal muscle: two functionally distinct actin-
RT regulatory proteins exhibit high structural homology.";
RL Biochemistry 29:7420-7425(1990).
RN [2]
RP STRUCTURE BY NMR, AND REVISIONS TO 53-54.
RX MEDLINE=21880669; Pubmed=11885570; DOI=10.1073/pnas.2422780866;
RA Bains N.P.S., Gorbatsyuk V.Y., Noworthy N.J., Robson S.A.,
RA Maciejewski M.W., dos Remedios C.G., King G.F.;
RT "Backbone and side-chain 1H, 15N, and 13C assignments for chick
RT cofilin.";
RL J. Biomol. NMR 22:193-194(2002).
```

CC -1- FUNCTION: Controls reversibly actin polymerization and depolymerization in a pH-sensitive manner. It has the ability to bind G- and F-actin in a 1:1 ratio of coflin to actin. It is the major component of intranuclear and cytoplasmic actin rods.

CC -1- SUBCELLULAR LOCATION: Intracellular and cytoplasmic; almost completely in nucleus in cells exposed to heat shock or 10% dimethyl sulfoxide.

CC -1- TISSUE SPECIFICITY: Widely distributed in various tissues.

CC -1- PTM: The phosphorylation of Ser-24 may prevent recognition of the nuclear localization signal.

CC -1- SIMILARITY: Belongs to the actin-binding proteins ADF family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: M55659; AA62732.1; -.

CC HSSP: P18282; 1AK6.

CC InterPro: IPR002108; Actbind_coflin.

CC Pfam: PF00241; Cofilin_ADF; 1.

CC PRINTS: PR00006; COFILIN.

CC ProDom: PD002129; Actbind_coflin; 1.

CC SMART: SM00102; ADF; 1.

CC PROSITE: PS00325; ACTIN DEPOLYMERIZING; 1.

CC Actin-binding; Cytoskeleton; Nuclear protein; Phosphorylation.

CC DOMAIN 30 34 Nuclear localization signal (Potential).

CC DOMAIN 106 125 Actin-binding (Potential).

CC MOD RES 24 24 Phosphoserine (By similarity).

CC SEQUENCE 166 AA; 18661 MW; 514685E940786EF0 CRC64;

Query Match Score 31; DB 1; Length 166;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 57 VGDIID 62

RESULT 16
Q6NZW3 PRELIMINARY; PRT; 166 AA.
AC Q6NZW3:
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE Cofilin 2 (Muscle).
GN Name=cf12;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; OC Cyprinidae; Danio.
OC NCB1_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.F.,
RA Jones S.J., Maitra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Director MGC Project;
RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC065947; AA65947.1; -.
DR HSSP: P60982; 1AK6
DR ZFIN: ZDB-GENE-040426-1815; cf12.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003779; F:actin binding; IEA.
DR InterPro: IPR002108; Actbind_coflin.
DR Pfam: PF00241; Cofilin_ADF; 1.
DR PRINTS: PR00006; COFILIN.
DR ProDom: PD002129; Actbind_coflin; 1.
DR SMART: SM00102; ADF; 1.
DR PROSITE: PS00325; ACTIN DEPOLYMERIZING; 1.
DR SEQUENCE 166 AA; 18568 MW; 795942E0F8F0A2D6 CRC64;

Query Match Score 31; DB 2; Length 166;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 57 VGDIID 62

RESULT 17
Q6S3Z5 PRELIMINARY; PRT; 180 AA.
AC Q6S3Z5:
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DE Gp171.
GN Name=171;
OS Mycobacteriophage Omega.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OC NCB1_TaxID=205879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2259260; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karchikyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Baratarov S.,
RA Krizkov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.,
RT "Origins of highly mosaic mycobacteriophage genomes."
RT Cell 113:171-182(2003).
DR EMBL: AY129338; AA12813.1; -.
DR SEQUENCE 180 AA; 20106 MW; C7396B1A20F0460 CRC64;

Query Match Score 31; DB 2; Length 180;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 94 IGDVGD 99

RESULT 18
Q9US43 PRELIMINARY; PRT; 189 AA.
AC Q9US43:

```

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Utracil phosphoribosyltransferase activity (Predicted); similar to S.
DE cerevisiae YHR128W.
GN Name=SPAC1002.17c;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN
RP
SEQUENCE FROM N.A.
RC STRAIN=972h-;
MEDLINE=21848401; Pubmed=11859360; DOI=10.1038/nature724;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,
RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odel C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,
RA Woodward J., Voickaert G., Aert R., Robben J., Grymoprez B.,
RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Muller-Auer S.,
RA Gabel C., Fuchs M., Dusterhoft A., Fritzc C., Holzer E., Moestl D.,
RA Hilbert H., Borzym K., Langer I., Beck A., Lehnart H., Reimhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leilaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rocher M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipakovski G.V., Tessey D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
DR EMBL; AL136078; CAB65617.1; -.
DR HSSP; Q26998; JULS.
DR GeneDB_Spombe; SPAC1002.17c; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0009116; F:nucleoside metabolism; IEA.
DR InterPro; IPR000508; Peptidase_526.
DR InterPro; IPR000836; PRTtransferase.
DR Pfam; PF00156; Ptdosyltransferase.
DR PROSITE; PS00501; SPASE_1.1; UNKNOWN_1.
DR GlycoSyLtransferase; Transferase.
SQ SEQUENCE 189 AA; 20476 MW; A1930220404EC933 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 189;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 177 VGDIGD 182

RESULT 19
084BH1 PRELIMINARY; PRT; 201 AA.
AC 084BH1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Via recombination cassette V1a13 (Fragment).
OS Borrelia afzelii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.

```

```

OX NCBI_TaxID=29518;
RN
RP
SEQUENCE FROM N.A.
RC STRAIN=ACAI;
RX MEDLINE=22493136; Pubmed=12603744;
RA Wang D., Botkin D.J., Norris S.J.;
RT "Characterization of the vis antigenic variation loci of the Lyme
RT disease spirochaetes Borrelia garinii ip90 and Borrelia afzelii
RT ACAI.";
RL Mol. Microbiol. 47:1407-1417(2003).
DR EMBL; AY106528; AAN87821.1; -.
DR HSSP; O68364; 1L8W.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000680; Borrelia_lipo.
DR Pfam; PF00921; Lipoprotein_2; 1.
KM Plasmid.
FT NON_TER 1
FT NON_TER 1
FT SEQUENCE 201 AA; 19522 MW; 174E3D9343708B01 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 201;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 34 IGDVGD 39

RESULT 20
09YB4 PRELIMINARY; PRT; 214 AA.
AC 09YB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coat protein.
OS Peanut stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=12313;
RN
RP
SEQUENCE FROM N.A.
RC STRAIN=western;
RX MEDLINE=97275907; Pubmed=9129668;
RA Hu C.-C., Aboul-Ata A.B., Naidu R.A., Ghaffari S.A.;
RT "Evidence for the occurrence of two distinct subgroups of peanut stunt
RT cucumovirus strains: molecular characterization of RNA3.";
RL J. Gen. Virol. 78:929-939(1997).
DR EMBL; U31366; AAC99857.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000247; Cucumovirus_coat.
DR Pfam; PF00760; Cucumo_coat; 1.
DR PRINTS; PR00222; CUCUMOCOAT.
DR ProDom; PD001284; Cucumovirus_coat; 1.
KM Coat protein.
SQ SEQUENCE 214 AA; 23765 MW; 9103CB8B8C0C9A93 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 214;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 173 VGDIGD 178

RESULT 21
019062 PRELIMINARY; PRT; 222 AA.
AC 019062;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)

```

DT 01-JAN-1998 (TREMELREL. 05, Last sequence update)
 DT 25-OCT-2004 (TREMELREL. 28, Last annotation update)
 DE C-reactive protein precursor.
 OS Sus scrofa (pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Landrace; TISSUE=Liver;
 RA Ozawa A., Matsunoto M., Kajikawa M., Hanazono M., Yasue H.;
 RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Schmedel S., Ponsukelli S., Schellander K., Wimmer K.;
 RT "Detection of SNPs and linkage and radiation hybrid mapping of the
 RT porcine C-reactive protein gene, CRP."
 RL Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB005545; BA021473.1; -;
 DR EMBL; AY714055; AAU13779.1; -;
 DR HSSP; P02741; 1809.
 DR InterPro; IPR008985; CONA_1ike_1ec_g1.
 DR InterPro; IPR001759; Pentaxin.
 DR Pfam; PF00354; Pentaxin; 1.
 DR PRINTS; PR00895; PENTAXIN.
 DR PRODOM; PD002153; Pentaxin; 1.
 DR SMART; SM00159; PTX; 1.
 DR PROSITE; PS00289; PENTAXIN; 1.
 KM Signal.
 FT CHAIN 1 19 Potential.
 FT SIGNAL 20 222 CRP mature peptide.
 SQ SEQUENCE 222 AA; 24924 MW; 451B7959D5FEB98 CRC64;
 Query Match 96.9%; Score 31; DB 2; Length 222;
 Best Local Similarity 83.3%; Pred. No. 3.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IGDIGD 6
 Db 169 VGDIGD 174
 RESULT 22
 CRP_CAVPO STANDARD; PRT; 225 AA.
 AC P49254;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE C-reactive protein precursor.
 GN Name=CRP; Synonyms=PTX1;
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
 CC NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Harley;
 RA MEDLINE=93252746; PubMed=8486600;
 RA Rubio N., Sharp P.M., Riles M., Zahedi K., Whitehead A.S.;
 RT "Structure, expression, and evolution of guinea pig serum amyloid P
 RT component and C-reactive protein."
 RL J. Biochem. 113:277-284(1993).
 CC -1- FUNCTION: Displays several functions associated with host defense:
 CC it promotes agglutination, bacterial capsular swelling,
 CC phagocytosis and complement fixation through its calcium-dependent
 CC binding to phosphorylcholine. Can interact with DNA and histones
 CC and may scavenge nuclear material released from damaged
 CC circulating cells (By similarity).
 CC -1- SUBUNIT: Homopentamer. Pentaxin (or pentraxin) have a discoid
 CC arrangement of 5 noncovalently bound subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Found in plasma.

CC -1- SIMILARITY: Belongs to the pentaxin family.
 CC -1- DATABASE: NAME=Protein Spotlight; NOTE=Issue 30 of January 2003;
 CC WWW="http://www.expasy.org/spotlight/articles/spl1030.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S60422; AAC60662.1; -;
 CC PIR; JX0259; JX0259.
 CC HSSP; P02741; 1809.
 CC InterPro; IPR008985; CONA_1ike_1ec_g1.
 CC InterPro; IPR001759; Pentaxin.
 CC Pfam; PF00354; Pentaxin; 1.
 CC PRINTS; PR00895; PENTAXIN.
 CC PRODOM; PD002153; Pentaxin; 1.
 CC SMART; SM00159; PTX; 1.
 CC PROSITE; PS00289; PENTAXIN; 1.
 KM Acute phase; Calcium; Pentaxin; Plasma; Signal.
 FT SIGNAL 1 19 By similarity.
 FT CHAIN 20 225 C-reactive protein.
 FT DOMAIN 20 225 Pentaxin.
 FT DISUFID 55 116 By similarity.
 SQ SEQUENCE 225 AA; 25225 MW; 47AF3E2DA60F2E CRC64;
 Query Match 96.9%; Score 31; DB 1; Length 225;
 Best Local Similarity 83.3%; Pred. No. 3.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IGDIGD 6
 Db 172 VGDIGD 177
 RESULT 23
 CRP_MESAU STANDARD; PRT; 225 AA.
 AC P49262;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE C-reactive protein precursor.
 GN Name=CRP; Synonyms=PTX1;
 OS Mesocricetus auratus (Golden hamster).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC Mesocricetus.
 CC NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=91369977; PubMed=1892852;
 RA Dowton S.B., Holden S.N.;
 RT "C-reactive protein (CRP) of the Syrian hamster."
 RL Biochemistry 30:9531-9538(1991).
 CC -1- FUNCTION: Displays several functions associated with host defense:
 CC it promotes agglutination, bacterial capsular swelling,
 CC phagocytosis and complement fixation through its calcium-dependent
 CC binding to phosphorylcholine. Can interact with DNA and histones
 CC and may scavenge nuclear material released from damaged
 CC circulating cells (By similarity).
 CC -1- SUBUNIT: Homopentamer. Pentaxin (or pentraxin) have a discoid
 CC arrangement of 5 noncovalently bound subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Found in plasma.
 CC -1- INDUCTION: By interleukin-1, interleukin-6, and TNF-alpha.
 CC -1- SIMILARITY: Belongs to the pentaxin family.
 CC -1- DATABASE: NAME=Protein Spotlight; NOTE=Issue 30 of January 2003;
 CC WWW="http://www.expasy.org/spotlight/articles/spl1030.html".
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; S56005; AAB1893.2; -.
 DR PIR; A40326; A40326.
 DR HSSP; P02741; 1LJ7.
 DR InterPro; IPR008985; ConA_like_1ec_g1.
 DR InterPro; IPR001759; Pentaxin.
 DR Pfam; PF00354; Pentaxin; 1.
 DR PRINTS; PR00895; PENTAXIN.
 DR ProDom; PD002153; Pentaxin; 1.
 DR SMART; SM00159; PTX; 1.
 DR PROSITE; PS00289; PENTAXIN; 1.
 DR Acute phase; Calcium; Pentaxin; Plasma; Signal.
 FT SIGNAL 1 19 By similarity.
 FT CHAIN 20 225 C-reactive protein.
 FT DOMAIN 20 225 Pentaxin.
 FT DISUFID 55 116 By similarity.
 SQ SEQUENCE 225 AA; 24822 MW; 00B439C45D58E05A CRC64;

Query Match 96.9%; Score 31; DB 1; Length 225;
 Best Local Similarity 83.3%; Pred. No. 3.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 Db 172 VGDIGD 177

RESULT 24
 CRP_MOUSE STANDARD; PRT; 225 AA.
 ID CRP_MOUSE
 AC P14847;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE C-reactive protein precursor.
 GN Name=Crp; Synonyms=Ptxl;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=69050112; PubMed=3190681;
 RA Ohnishi S., Maeda S., Nishiguchi S., Arao T., Shimada K.;
 RT "Structure of the mouse C-reactive protein gene."
 RL Biochem. Biophys. Res. Commun. 156:814-822(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBA/J; TISSUE=Liver;
 RX MEDLINE=90179719; PubMed=2310378;
 RA Whitehead A.S.; Zahedi K.; Rits M.; Mortensen R.F.; Lelias J.M.;
 RT "Mouse C-reactive protein. Generation of cDNA clones, structural
 RL analysis, and induction of mRNA during inflammation."
 CC -1- FUNCTION: Displays several functions associated with host defense:
 CC it promotes agglutination, bacterial capsular swelling,
 CC phagocytosis and complement fixation through its calcium-dependent
 CC binding to phosphorylcholine. Can interact with DNA and histones
 CC and may scavenge nuclear material released from damaged
 CC circulating cells.
 CC -1- SUBUNIT: Homopentamer. Pentaxin (or pentraxin) have a discoid
 CC arrangement of 5 noncovalently bound subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Found in plasma.
 CC -1- INDUCTION: The concentration of CRP in plasma increases greatly
 CC during acute phase response to tissue injury, infection or other

CC inflammatory stimuli.
 CC -1- SIMILARITY: Belongs to the pentaxin family.
 CC -1- DATABASE: NAME=Protein Spotlight; NOTE=Issue 30 of January 2003;
 CC WWW=<http://www.ebi.ac.uk/spotlight/articles/spot1030.html>.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; X13588; CA31928.1; -.
 DR EMBL; X17496; CA35531.1; -.
 DR PIR; A31583; A31583.
 DR HSSP; P02741; 1LJ7.
 DR MGD; MGI:88512; Crp.
 DR InterPro; IPR008985; ConA_like_1ec_g1.
 DR InterPro; IPR001759; Pentaxin.
 DR Pfam; PF00354; Pentaxin; 1.
 DR PRINTS; PR00895; PENTAXIN.
 DR ProDom; PD002153; Pentaxin; 1.
 DR SMART; SM00159; PTX; 1.
 DR PROSITE; PS00289; PENTAXIN; 1.
 DR Acute phase; Calcium; Pentaxin; Plasma; Signal.
 FT SIGNAL 1 19 C-reactive protein.
 FT CHAIN 20 225 Pentaxin.
 FT DOMAIN 20 225 Pentaxin.
 FT DISUFID 55 116 By similarity.
 FT CONFLICT 134 134 A -> P (in Ref. 2).
 SQ SEQUENCE 225 AA; 25361 MW; 15586785E46043D9 CRC64;

Query Match 96.9%; Score 31; DB 1; Length 225;
 Best Local Similarity 83.3%; Pred. No. 3.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 Db 172 VGDIGD 177

RESULT 25
 Q91XB3 PRELIMINARY; PRT; 225 AA.
 ID Q91XB3
 AC Q91XB3;
 DT 01-DEC-2001 (TREMURel. 19, Created)
 DT 01-DEC-2001 (TREMURel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMURel. 25, Last annotation update)
 DE C-reactive protein, pentaxin related.
 GN Name=Crp;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Liver;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
 RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
 RA Altshul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
 RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Heif F.;
 RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
 RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
 RA Brownstein M.J.; Udén T.B.; Toshiyuki S.; Carninci P.; Prange C.;
 RA Rana S.S.; Loughlano N.A.; Peters G.J.; Abramson R.D.; Mullaney S.J.;
 RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
 RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
 RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
 RA Fahey J.; Helton E.; Kettelman M.; Madan A.C.; Rodrigues S.; Sanchez A.;
 RA Whitting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
 RA Blakeley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;

Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S., Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E., Jones S.J., Mair M.A., "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

[2]
RN SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC011124; AAH1124.1; -
DR HSSP: P02741; 1LJ7.
DR MGD: MGI:86512; Crp.
DR GO: GO:0005615; Cytoplasmic space; TAS.
DR Pfam: PF00354; Pentaxin; 1.
DR PRINTS: PR00895; PENTAXIN.
DR PRODOM: PD002153; Pentaxin; 1.
DR SMART: SM00159; PTX; 1.
DR PROSITE: PS00289; PENTAXIN; 1.
DR SEQUENCE 225 AA; 25387 MW; 0B5FD795B57052D8 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 225;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 172 VGDIGD 177

RESULT 26
CRP RAT STANDARD; PRT; 230 AA.
AC P48199;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE C-reactive protein precursor.
GN Name=Crp; Synonyms=Pcx1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=92147634; PubMed=1737750;
RA Nassouli M., Sambasivam H., Azadi P., Dell A., Morris H.R., Nagpurkar A., Mockery S., Murray R.K.;
RT "Derivation of the amino acid sequence of rat C-reactive protein from cDNA cloning with additional studies on the nature of its dimeric component";
RT J. Biol. Chem. 267:2947-2954 (1992).

-1- FUNCTION: Displays several functions associated with host defense: it promotes agglutination, bacterial capsular swelling, phagocytosis and complement fixation through its calcium-dependent binding to phosphorylcholine. Can interact with DNA and histones and may scavenge nuclear material released from damaged circulating cells (by similarity).
-1- SUBUNIT: Homopentamer. Pentaxin (or pentaxin) have a discoid arrangement of 5 noncovalently bound subunits. Two of the five chains form a dimer linked by two interchain disulfide bonds located in the C-terminal heptapeptide and specific to rat Crp.
-1- SUBCELLULAR LOCATION: Secreted.
-1- TISSUE SPECIFICITY: Found in plasma.
-1- PTM: The last two cysteines are involved either in interchain disulfide bonds or in an intrachain bond.
-1- SIMILARITY: Belongs to the pentaxin family.
-1- DATABASE: NAME=Protein Spotlight; NOTE=Issue 30 of January 2003; WWW="http://www.expasy.org/spotlight/articles/spcl030.html".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
DR EMBL: M83176; AAA0964.1; -
DR PIR: A42579; A42579.
DR HSSP: P02741; 1LJ7.
DR GLYCOSULEDB: P48199; -
DR InterPro: IPR008985; ConA_like_1ec_g1.
DR InterPro: IPR001759; Pentaxin.
DR Pfam: PF00354; Pentaxin; 1.
DR PRINTS: PR00895; PENTAXIN.
DR PRODOM: PD002153; Pentaxin; 1.
DR SMART: SM00159; PTX; 1.
DR PROSITE: PS00289; PENTAXIN; 1.
DR Acute phase; Calcium; Direct protein sequencing; Glycoprotein;
KW Pentaxin; Plasma; Signal.
FT SIGNAL 1 19 By similarity.
FT CHAIN 20 230 C-reactive protein.
FT DOMAIN 20 230 Pentaxin.
FT DISULFID 55 114 Interchain (in linked chains).
FT DISULFID 227 227 Interchain (in linked chains).
FT DISULFID 228 228 In unlinked chains.
FT DISULFID 227 228 N-linked (GlcNAc...).
FT CARBOHYD 147 147 /FTID=CAR_000154.
FT FT D8CF6BF72376309 CRC64;
SQ SEQUENCE 230 AA; 25468 MW; 8E2C3D2ED9AEEAAD CRC64;

Query Match 96.9%; Score 31; DB 1; Length 230;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 170 VGDIGD 175

RESULT 27
O8GHWO PRELIMINARY; PRT; 247 AA.
ID O8GHWO;
AC O8GHWO;
DT 01-MAR-2003 (TRENDEL. 23, Created)
DT 01-MAR-2003 (TRENDEL. 23, Last sequence update)
DT 01-JUN-2003 (TRENDEL. 24, Last annotation update)
DE Hypothetical protein O8F89.
GN Name=O8F89;
OS Pseudomonas resinovorans.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OC NCBI_TaxID=53412;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22435905; PubMed=12547188; DOI=10.1016/S0022-2836(02)01400-6;
RA Maeda K., Nojiri H., Shintani M., Yoshida T., Haba H., Omori T.;
RT "Complete nucleotide sequence of carbazole/dioxin-degrading plasmid pCAR1 in Pseudomonas resinovorans strain CA10 indicates its mosaicity RT and the presence of large catabolic transposon Tn4676";
RT J. Mol. Biol. 326:21-33 (2003).
RL T. Mol. Biol. 326:21-33 (2003).
DR EMBL: AB088420; BAC41609.1;
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 247 AA; 28348 MW; 8E2C3D2ED9AEEAAD CRC64;

Query Match 96.9%; Score 31; DB 2; Length 247;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 70 VGDIGD 75

```

RESULT 28
084176 PRELIMINARY; PRT; 252 AA.
ID 084176
AC 084176;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative short-chain dehydrogenase.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Wautersia.
NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=112R4;
RX MEDLINE=22445581; PubMed=12557390;
RA Wang Y., Zhang Y., Ding J., Liu Y., Wang J., Yu Z.;
RT "Cloning, sequence analysis of imidase gene from Alcaligenes eutrophus
RT and its expression in E. coli.";
RL Wei Sheng Wu Xue Bao 42:153-162(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=112R4;
RA Zhang Y.-Z., Yu W., Ding J.-Y., Yu Z.-H., Liu Y.-J., Wang J.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
DR EMBL; AF373287; AA042633.1; -.
DR HSSP; P97852; IG26.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
KM Oxidoreductase.
SQ SEQUENCE 252 AA; 26128 MW; 560222B38920826 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 252;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 64 IGDVGD 69

RESULT 29
041345 PRELIMINARY; PRT; 281 AA.
ID 041345
AC 041345;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE STA1-12.
GN Name=STA1-12;
OS Silene pratensis (White campion) (Lychnis alba).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Silene.
NCBI_TaxID=52853;
RN [1]
RP SEQUENCE FROM N.A.
RC Hardack S., De Y., Koutnikova H., Saeidler H., Grant S.R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
DR EMBL; U53828; AAB42054.1; -.
DR HSSP; P25716; 1Q7B.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KM Oxidoreductase.
SQ SEQUENCE 281 AA; 28969 MW; 0A54D7DB8FCE9CA4 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 281;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 213 VGDIGD 218

RESULT 30
042407 PRELIMINARY; PRT; 281 AA.
ID 042407
AC 042407;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE STA1-18 (STA1-2).
GN Name=STA1-18; Synonyms=STA1-2;
OS Silene pratensis (White campion) (Lychnis alba).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Silene.
NCBI_TaxID=52853;
RN [1]
RP SEQUENCE FROM N.A.
RC Hardack S., De Y., Koutnikova H., Saeidler H., Grant S.R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
DR EMBL; U53829; AAB42055.1; -.
DR EMBL; U53827; AAB42053.1; -.
DR HSSP; P25716; 1Q7B.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KM Oxidoreductase.
SQ SEQUENCE 281 AA; 28969 MW; 0A54D7DB8FCE9CA4 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 281;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 213 VGDIGD 218

RESULT 31
CCDS_CABEL STANDARD; PRT; 284 AA.
ID CCDS_CABEL
AC P91285;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cuticle collagen dpy-5 (Dumpy protein 5).
GN Name=dpy-5; ORFNames=F27C1.8;
OS Caenorhabditis elegans.

```

```

DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KM Oxidoreductase.
SQ SEQUENCE 281 AA; 28940 MW; 0E921AD881A2AC4 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 281;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 213 VGDIGD 218

RESULT 30
042407 PRELIMINARY; PRT; 281 AA.
ID 042407
AC 042407;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE STA1-18 (STA1-2).
GN Name=STA1-18; Synonyms=STA1-2;
OS Silene pratensis (White campion) (Lychnis alba).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Silene.
NCBI_TaxID=52853;
RN [1]
RP SEQUENCE FROM N.A.
RC Hardack S., De Y., Koutnikova H., Saeidler H., Grant S.R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
DR EMBL; U53829; AAB42055.1; -.
DR EMBL; U53827; AAB42053.1; -.
DR HSSP; P25716; 1Q7B.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KM Oxidoreductase.
SQ SEQUENCE 281 AA; 28969 MW; 0A54D7DB8FCE9CA4 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 281;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
Db 213 VGDIGD 218

RESULT 31
CCDS_CABEL STANDARD; PRT; 284 AA.
ID CCDS_CABEL
AC P91285;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cuticle collagen dpy-5 (Dumpy protein 5).
GN Name=dpy-5; ORFNames=F27C1.8;
OS Caenorhabditis elegans.

```



```

CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology."
RN Science 282:2012-2018(1998).
[2]
RP IDENTIFICATION.
RA Thacker C.M., Rose A.M.;
RT "dpy-5 encodes a cuticle collagen.";
RL (ex) Worm Breeder's Gazette 15(1):54(1997).
CC -1- FUNCTION: Nematode cuticles are composed largely of collagen-like
CC proteins. The cuticle functions both as an exoskeleton and as a
CC barrier to protect the worm from its environment.
CC -1- SUBUNIT: Collagen polypeptide chains are complexed within the
CC cuticle by disulfide bonds and other types of covalent cross-links
CC (By similarity).
CC -1- PTM: May be a substrate of bli-4.
CC -1- SIMILARITY: Belongs to the cuticular collagen family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL; U80441; AAB37656.1; -.
DR PIR; T29528; T29528.
DR WormBase; WBGene00001067; dpy-5.
DR WormPeP; F27C1.8; CE09720.
DR InterPro; IPR008161; C1g_helix.
DR InterPro; IPR002486; Col_cuticle_N.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01484; Col_cuticle_N; 1.
DR Pfam; PF01391; Collagen; 2.
DR ProDom; PD000007; C1g_helix; 2.
DR KEGG; collagen; Cuticle; Multigene family; Repeat; Structural protein.
DR KEGG; collagen; Cuticle; Multigene family; Repeat; Structural protein.
FT DOMAIN 94 126
FT DOMAIN 143 270
FT DOMAIN 143 270
SQ SEQUENCE 264 AA; 28216 MW; 8BED98AED554B454 CRC64;
Query Match 96.9%; Score 31; DB 1; Length 284;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
DB 178 VSDIGD 183

```

```

RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Matanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
RA Miyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus Bp-1."
RN DNA Res. 9:123-130(2002).
CC -1- SIMILARITY: Belongs to the sugar epimerase family.
DR EMBL; AP005376; BAC09758.1; -.
DR HSPB; Q8T8E9; 1G78.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GALE.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; gale; 1.
KW Complete proteome; NAD.
SQ SEQUENCE 308 AA; 34440 MW; 1DF4B24A4A31E2FC2 CRC64;
Query Match 96.9%; Score 31; DB 2; Length 308;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
DB 25 VSDIGD 30

```

RESULT 33

```

ID Q67KV0 PRELIMINARY; PRT; 328 AA.
AC Q67KV0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE UDP-glucose 4-epimerase.
GN ORFNames=STH2711;
OS Synbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM4863;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium
RT thermophilum."
RT Submitted (Apr-2004) to the EMBL/GenBank/DDBJ databases.
CC -1- SIMILARITY: Belongs to the sugar epimerase family.
DR EMBL; AP006840; BAD41696.1; -.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GALE.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; gale; 1.
KW NAD.
SQ SEQUENCE 328 AA; 36120 MW; 5E07FACDE4D889A3 CRC64;
Query Match 96.9%; Score 31; DB 2; Length 328;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 IGDIGD 6
DB 44 IGDVGD 49

```

RESULT 34

```

ID Q82GL3 PRELIMINARY; PRT; 328 AA.
AC Q82GL3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

```

DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Putative integral membrane protein.
GN OrderedLocustNames=SAV3884;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL: AP005036; BAC71596.1; -.
DR InterPro: IPR011044; Amine_DH_B_like.
KW Complete proteome.
SQ SEQUENCE 328 AA; 34184 MW; 7E08F2068F22456 CRC64;

Query Match          96.9%; Score 31; DB 2; Length 328;
Best Local Similarity 83.3%; Pred. No. 5,5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
Db      105 VGDIGD 110

RESULT 35
Q9FZ20 PRELIMINARY; PRT; 348 AA.
ID Q9FZ20;
AC Q9FZ20;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Putative integral membrane protein.
GN ORFNames=SCD19.12;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/41714a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.B., Quail M.A., Kessler H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kisser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabhinowitsch B., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrett B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL399119; CAC08313.1; -.
DR InterPro: IPR011044; Amine_DH_B_like.

```

```

KW Complete proteome.
SQ SEQUENCE 348 AA; 36391 MW; 1F78E19CE1AFB75 CRC64;

Query Match          96.9%; Score 31; DB 2; Length 348;
Best Local Similarity 83.3%; Pred. No. 5,9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
Db      128 VGDIGD 133

RESULT 36
Q9RZB3 PRELIMINARY; PRT; 364 AA.
ID Q9RZB3;
AC Q9RZB3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DE 01-MAY-2004 (TrEMBLrel. 26, last annotation update)
DE Thymidine diphosphoglucose 4,6-dehydratase.
GN OrderedLocustNames=DRA0041;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Bisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Halt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA McFie K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -1- SIMILARITY: Belongs to the sugar epimerase family.
DR EMBL: AE001862; AAF12268.1; -.
DR PIR: C75597; C75597.
DR HSP; P27830; IBXK.
DR TIGR; DRA0041; -.
DR GO: GO:0008460; F:GTP-glucose 4,6-dehydratase activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro: IPR002347; Adh_short_C2.
DR InterPro: IPR005888; dTDP_gluC_dehyd.
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase_1.
DR PRINTS: PRO0081; GDMRDH.
DR TIGRPFAMs: TIGR01181; dTDP_gluC_dehyd; 1.
KW Complete proteome; NAD.
SQ SEQUENCE 364 AA; 40885 MW; A3C2F6311AF469EC CRC64;

Query Match          96.9%; Score 31; DB 2; Length 364;
Best Local Similarity 83.3%; Pred. No. 6,1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IGDIGD 6
Db      70 VGDIGD 75

RESULT 37
Q7XTS7 PRELIMINARY; PRT; 588 AA.
ID Q7XTS7;
AC Q7XTS7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)

```

```

DE OSJNB0008M17.6 protein.
GN Name=OSJNB0008M17.6;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzae; Oryza.
NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Yang K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Yu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL662950; CAD41791.2; -.
DR Gramene; Q7XTS7; -.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR InterPro; IPR001878; Znf_CCHC.
DR SMART; SM00343; Znf_C2HC; 1.
DR SEQUENCE 588 AA; 64054 MW; 191731C6B1883E8A CRC64;

Query Match 96.9%; Score 31; DB 2; Length 588;
Best Local Similarity 83.3%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 313 VGDI GD 318

RESULT 38
Q98KR2 PRELIMINARY; PRT; 693 AA.
ID Q98KR2;
AC Q98KR2;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE NADH-ubiquinone dehydrogenase chain 3.
GN OrderedLocustNames=ml11362;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobium; Proteobacteria; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAP303099;
RC MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matano A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
CC DNA Rec. 7:331-338(2000).
CC -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -1- SIMILARITY: Belongs to the complex I 75 kDa subunit family.
CC EMBL; AP002997; BAB48752.1; -.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0005506; P:iron ion binding; IEA.
DR GO; GO:0008137; P:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016651; P:oxidoreductase activity, acting on NADH or . . .; IEA.

```

```

DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR000283; ComplexI_75k.
DR InterPro; IPR001041; Ferredoxin.
DR InterPro; IPR006656; Molybdopterin.
DR InterPro; IPR010228; NuoG.
DR Pfam; PF00111; Fer2; 1.
DR Pfam; PF00384; Molybdopterin; 1.
DR TIGRFAMs; TIGR01973; NuoG; 1.
DR PROSITE; PS00642; COMPLEXI_75K_2; 1.
DR PROSITE; PS00643; COMPLEXI_75K_3; 1.
DR 2fe-2s; 4fe-4s; Complete proteome; Iron; Iron-sulfur; Metal-binding;
KW NAD; Oxidoreductase; Quinone; Ubiquinone.
SQ SEQUENCE 693 AA; 74544 MW; 6CC39A61B38A7631 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 693;
Best Local Similarity 83.3%; Pred. No. 1.2e+03; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
Db 409 IGDIGD 414

RESULT 39
GYRB_RICCN STANDARD; PRT; 807 AA.
ID GYRB_RICCN
AC Q92H87;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA gyrase subunit B (EC 5.99.1.3).
GN Name=gyrB; OrderedLocustNames=RC0884;
OS Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Bacteria; Proteobacteria; Rickettsiales; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RC MEDLINE=21442074; PubMed=11557893; DOI=10.1126/science.1061471;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cosset P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."
RL Science 293:2093-2098(2001).
CC -1- FUNCTION: DNA gyrase negatively supercoils closed circular double-
CC stranded DNA in an ATP-dependent manner and also catalyzes the
CC interconversion of other topological isomers of double-stranded
CC DNA rings, including catenanes and knotted rings.
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
CC enzyme forms an A2B2 tetramer.
CC -1- SIMILARITY: Belongs to the type II topoisomerase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB008642; AAL03422.1; -.
CC PIR; D97810; D97810.
CC HSP; P06982; LA76.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR002288; DNA_gyraseB_C.
DR InterPro; IPR000565; DNA_gyraseB.
DR InterPro; IPR001241; DNA_topoisomII.
DR InterPro; IPR006171; Toprim_dom.

```

DR Pfam; PF00204; DNA_gyraseB; 1.
 DR Pfam; PF00986; DNA_gyraseC; 1.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF01751; Toprim; 1.
 DR PRINTS; PR00418; TP12FAMILY.
 DR ProDom; PD149633; DNA_gyrase_B; 1.
 DR SMART; SM00387; HATPase_C; 1.
 DR SMART; SM00433; TOP2C; 1.
 DR TIGRFAMs; TIGR01059; gyrb; 1.
 DR PROSITE; PS00177; TOPOISOMERASE_II; 1.
 KM ATP-binding; Complete proteome; Isomerase; Topoisomerase.
 SQ SEQUENCE 807 AA; 90794 MW; 12814B43F12132F9 CRC64;

Query Match 96.9%; Score 31; DB 1; Length 807;
 Best Local Similarity 83.3%; Pred. No. 1.4e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
 Db 35 IGDVGD 40

RESULT 40
 GYRB_RICPR STANDARD; PRT; 807 AA.
 ID GYRB_RICPR
 AC Q9ZCX2;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE DNA gyrase subunit B (EC 5.99.1.3).
 GN Name=gyrB; OrderedLocNames=RP580;
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 NCBI_TaxID=782;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893; DOI=10.1038/24094;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Scheraga H., Altmann U.C.M., Podowski R.M., Naslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria."
 RL Nature 396:133-140(1998).
 CC -1- FUNCTION: DNA gyrase negatively supercoils closed circular double-
 CC stranded DNA in an ATP-dependent manner and also catalyzes the
 CC interconversion of other topological isomers of double-stranded
 CC DNA rings, including catenanes and knotted rings.
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 CC of double-stranded DNA.
 CC -1- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
 CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
 CC enzyme forms an A2B2 tetramer.
 CC -1- SIMILARITY: Belongs to the type II topoisomerase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

DR Pfam; PF00986; DNA_gyraseB_C; 1.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF01751; Toprim; 1.
 DR PRINTS; PR00418; TP12FAMILY.
 DR ProDom; PD149633; DNA_gyrase_B; 1.
 DR SMART; SM00387; HATPase_C; 1.
 DR SMART; SM00433; TOP2C; 1.
 DR TIGRFAMs; TIGR01059; gyrb; 1.
 DR PROSITE; PS00177; TOPOISOMERASE_II; 1.
 KM ATP-binding; Complete proteome; Isomerase; Topoisomerase.
 SQ SEQUENCE 807 AA; 90823 MW; 5DEB5ABBD1F17D1 CRC64;

Query Match 96.9%; Score 31; DB 1; Length 807;
 Best Local Similarity 83.3%; Pred. No. 1.4e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
 Db 35 IGDVGD 40

RESULT 41
 GYRB_RICPR PRELIMINARY; PRT; 807 AA.
 ID GYRB_RICPR
 AC Q7P921;
 DT 01-MAR-2004 (TIGR01059, Created)
 DT 01-MAR-2004 (TIGR01059, Last sequence update)
 DT 01-MAR-2004 (TIGR01059, Last annotation update)
 DE DNA gyrase subunit B.
 GN Name=gyrB; OrderedLocNames=RP580;
 OS Rickettsia sibirica.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 NCBI_TaxID=35793;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Malek J.A., Bremeyer M.E., Dasch G.A.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DNA gyrase negatively supercoils closed circular double-
 CC stranded DNA in an ATP-dependent manner and also catalyzes the
 CC interconversion of other topological isomers of double-stranded
 CC DNA rings, including catenanes and knotted rings (by similarity).
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 CC of double-stranded DNA.
 CC -1- SIMILARITY: Belongs to the type II topoisomerase family.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AABW0100001; EAA26372.1; -.
 DR HSSP; P06982; 1AJ6.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
 DR GO; GO:0006304; P:DNA modification; IEA.
 DR GO; GO:0006265; P:DNA topological change; IEA.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR002288; DNA_gyrase_C.
 DR InterPro; IPR011558; DNA_gyrase_B.
 DR InterPro; IPR001241; DNA_topoisomII.
 DR InterPro; IPR011557; Gyrb.
 DR InterPro; IPR006171; Toprim_dom.
 DR Pfam; PF00204; DNA_gyraseB; 1.
 DR Pfam; PF00986; DNA_gyraseC; 1.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF01751; Toprim; 1.
 DR PRINTS; PR00418; TP12FAMILY.
 DR ProDom; PD149633; DNA_gyrase_B; 1.
 DR TIGRFAMs; TIGR01059; gyrb; 1.
 DR PROSITE; PS00177; TOPOISOMERASE_II; 1.
 KM ATP-binding; Isomerase; Topoisomerase.
 SQ SEQUENCE 807 AA; 90651 MW; 53FED1A519AF358A CRC64;

Query Match 96.9%; Score 31; DB 2; Length 807;
 Best Local Similarity 83.3%; Pred. No. 1.4e+03;

	Matches	5; Conservative	1; Mismatches	0; Indels	0; Gaps
QY	1	IGDIGD	6		
			:		
Db	35	IGDVGD	40		

QY 1 IGDIGD 6
:|||||
Db 788 VGDIGD 793

RESULT 44

Q8C777 PRELIMINARY; PRT; 1224 AA.

08C777; 08C777;
DT 01-MAR-2003 (TEMBRel. 23, Created)
DT 01-MAR-2003 (TEMBRel. 25, Last sequence update)
DT 01-OCT-2003 (TEMBRel. 25, Last annotation update)
DE Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched
DE library, clone:DJ30046B13 product:mitochondosome maintenance deficient
DE (S. cerevisiae) 3-associated protein, full insert sequence.
DE (Fragment).
GN Name=Mcm3ap;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).
RL Nature 409:685-690(2001).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
[3]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."; Nature 420:563-573(2002).
[4]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630(2000).
[5]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
Kono H., Akiyama T., Nishi K., Kikunishi T., Tashiro H., Itoh M.,
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaishuwa K.,
Fujisawa S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system: 384-format sequencing pipeline with 384 multiplexed sequencer."; Genome Res. 10:1757-1771(2000).
[6]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,

RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Koda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK052382; BAC34967.1; -.
DR MGI; MGI:1930089; Mcm3ap.
DR GO; GO:0005737; C:cycloplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR005062; SAC3_GANP.
DR Pfam; PF03399; SAC3_GANP; 1.
FT NON_TER 1
FT TER 1224
SQ SEQUENCE 1224 AA; 136841 MW; 25463FC6CDB03A67 CRC64;

Query Match 96.3%; Score 31; DB 2; Length 1224;
Best Local Similarity 83.3%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
:|||||
Db 790 VGDIGD 795

RESULT 45

Q7USG5 PRELIMINARY; PRT; 1637 AA.

07USG5; ID Q7USG5;
DT 01-OCT-2003 (TEMBRel. 25, Created)
DT 01-OCT-2003 (TEMBRel. 25, Last sequence update)
DT 01-MAR-2004 (TEMBRel. 26, Last annotation update)
DE Sensory transduction histidine kinase.
GN OrderedLocustNames=RB4511;
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schleener H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp. strain 1."; Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; BX294140; CAD73831.1; -.
DR HSP; Q9ZWC6; JUNT.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. .; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR000673; CheB_methylase.
DR InterPro; IPR000780; CheR_methylase.
DR InterPro; IPR011066; CheY_like.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR InterPro; IPR002114; Hpr_Ser_S.
DR InterPro; IPR002052; Hpr_Mtase.

```

DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR000700; PAS-associ C.
DR InterPro; IPR001789; Response_Reg.
DR Pfam; PF01333; Cheb_methylase; 1.
DR Pfam; PF01739; Cher; 1.
DR Pfam; PF03705; Cher_N; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00512; HSKA; 1.
DR Pfam; PF00785; PAC; 3.
DR Pfam; PF00989; PAC; 4.
DR Pfam; PF00722; Response_reg; 1.
DR Pfam; PF005328; Cheb_methylase; 1.
DR Pfam; PF000039; Response_reg; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HSKA; 1.
DR SMART; SM00138; Metrc; 1.
DR SMART; SM00066; PAC; 3.
DR SMART; SM00091; PAS; 4.
DR SMART; SM00448; REC; 1.
DR TIGRPFAMs; TIGR00229; sensory_box; 3.
DR PROSITE; PS50122; CHEB; 1.
DR PROSITE; PS50123; CHEB; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS00092; NG_MTPASE; UNKNOWN_1.
DR PROSITE; PS50113; PAC; 2.
DR PROSITE; PS50112; PAS; 3.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW TRANSFERASE; 1637 AA; 183571 MW; 20EB2A0549553F49 CRC64;
SQ SEQUENCE

```

```

Query Match          96.9%; Score 31; DB 2; Length 1637;
Best Local Similarity 83.3%; Pred. No. 2.9e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 IGDIGD 6
Db 472 VGDIGD 477

```

```

RESULT 46
O6NMV4 PRELIMINARY; PRT; 1732 AA.
AC O6NMV4;
DT 05-JUL-2004 (TRMBLrel. 27, Created)
DT 05-JUL-2004 (TRMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRMBLrel. 27, Last annotation update)
DE Mcm3ap protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=2238857; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Mariani K., Farmer A.A., Rubin G.M., Hong L.,
RA Starletton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mallyah S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,
RA Villalón D.K., Muny D.M., Sodegryn E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton B., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```

```

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Kirywiniski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Struhsberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC067414; AAH67414.1; -
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR005062; SAC3_GANP.
DR Pfam; PF03399; SAC3_GANP; 1.
DR NON_TER
SQ SEQUENCE 1732 AA; 192838 MW; 8CC0CB86FDF6EDC3 CRC64;

```

```

Query Match          96.9%; Score 31; DB 2; Length 1732;
Best Local Similarity 83.3%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 IGDIGD 6
Db 1149 VGDIGD 1154

```

```

RESULT 47
M3A_MOUSE STANDARD; PRT; 1971 AA.
ID M3A_MOUSE
AC O9WU9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 80 kDa MCM3-associated protein (GANP protein).
GN Name=Mcm3ap; Synonyms=GANP, Map80;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/C X NZB;
RX MEDLINE=20197882; PubMed=10733502;
RA Kuwahara K., Yoshida M., Komodo E., Sakata A., Watanabe Y., Abe E.,
RA Kouno Y., Tomiyasu S., Fujimura S., Tokuhisa T., Kimura H., Ezaki T.,
RA Sakaguchi N.;
RT "A novel nuclear phosphoprotein, GANP, is up-regulated in centrocytes
RT of the germinal center and associated with MCM3, a protein essential
RT for DNA replication."
RT Blood 95:2321-2328(2000).
CC -1- FUNCTION: May be involved in the nuclear localization pathway of
CC MCM3 (By similarity).
CC -1- SUBUNIT: Interacts with MCM3.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -1- PTM: Phosphorylated (Probable).
CC -1- SIMILARITY: Belongs to the SAC3 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AJ006590; CAB4241.1; -
DR MGD; MGI:1330089; Mcm3ap.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR005062; SAC3_GANP.
DR Pfam; PF03399; SAC3_GANP; 1.

```

KW Nuclear protein; Phosphorylation.
 SQ SEQUENCE 1971 AA; 217138 MW; 5F342E256C007E24 CRC64;
 Query Match 96.9%; Score 31; DB 1; Length 1971;
 Best Local Similarity 83.3%; Pred. No. 3.5e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
 Db 1388 VGDIGD 1393

RESULT 48
 Q7TS87 PRELIMINARY; PRT; 1971 AA.
 ID Q7TS87
 AC Q7TS87;
 DT 01-JUN-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mitochondrion maintenance protein 3 associated protein.
 DE Name=Mcm3ap;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC052452; AAH52452.1; -
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR005062; SAC3_GAMP.
 DR Pfam; PF03399; SAC3_GAMP; 1.
 SQ SEQUENCE 1971 AA; 217361 MW; FA9CDB20D1441B02 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 1971;
 Best Local Similarity 83.3%; Pred. No. 3.5e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
 Db 1388 VGDIGD 1393

RESULT 49
 Q80TK5 PRELIMINARY; PRT; 1992 AA.
 ID Q80TK5

AC Q80TK5;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE KIAA0572 protein (Fragment).
 DE KIAA0572 protein (Fragment).
 GN Name=Mcm3ap; Synonyms=MKIAA0572;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yusa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 DR EMBL; AK122313; BAC65595.1; -
 DR MGD; MGI:1930089; Mcm3ap.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR005062; SAC3_GAMP.
 DR Pfam; PF03399; SAC3_GAMP; 1.
 FT NON TER 1
 SQ SEQUENCE 1992 AA; 219884 MW; C980215AE6B2D440 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 1992;
 Best Local Similarity 83.3%; Pred. No. 3.5e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGDIGD 6
 :|||||
 Db 1409 VGDIGD 1414

RESULT 50
 Q7RHK7 PRELIMINARY; PRT; 2251 AA.
 ID Q7RHK7
 AC Q7RHK7;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=PY01809;
 OS Plasmodium yoelii yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NC NCBI_TaxID=73239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17XN1;
 RX PubMed=13368865; DOI=10.1038/nature01099;
 RA Carlson J.M., Anguilo S.V., Suh B.B., Kooij T.W., Perteu M.,
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Seleguot J.D., Koo H.L.,
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
 RA Shallow S.J., van Aken S.B., Riedmiller S.B., Feldblum T.V.,
 RA Cho J.K., Quackenbush J., Sedegah M., Shoab A., Cummings L.M.,
 RA Florens L., Yates F.R. II, Raine J.D., Sinden R.E., Harris W.A.,
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B.,
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
 RA Carucci D.J.;
 RT "Genome sequence and comparative analysis of the model rodent malaria
 RT parasite Plasmodium yoelii yoelii.";
 RL Nature 419:512-519(2002).
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAB01000489; EAA21180.1; -
 DR InterPro; IPR001623; DnaU_N.

DR InterPro; IPR006209; EGF-like.
DR PROSITE; PS00636; DNAJ_1; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 2251 AA; 267817 MW; F3B6FCB24BBD5492 CRC64;

Query Match 96.9%; Score 31; DB 2; Length 2251;
Best Local Similarity 83.3%; Pred. No. 4e+03; 0; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGDIGD 6
|||:
Db 806 IGDVGD 811

Search completed: July 27, 2005, 01:20:29
Job time : 178 secs

This Page Blank (uspto)